

17 April 1980

Conserve energy, conserve cash

"YOU can't run Britain like a grocer's shop!" said an exasperated trade union leader recently. He was referring not only to Mrs Thatcher's humble economic education, but also to the government's uncompromising arithmetic policy of restricting current expenditure by taking pennies, regardless, off every item.

The only lines in Britain's shop to remain untouched have been ones close to Mrs Thatcher's heart: defence, law and order, and nuclear power — the latter benefiting from a £10 billion spending programme over the next decade. Energy conservation, on the other hand, she has been content to leave to market forces.

Officials in the Department of Energy are not impressed, according to an internal memo recently leaked to the media, which draws attention to "a sharp and visible contraction in the government's conservation programme". The problem, says the memo, may be that "apart from improving the building regulations, we have in fact no proposals on the conservation front which do not call for extra resources either in terms of money or staff. At the present time, both are clearly not starters."

Another leaked document — this time minutes of a meeting of the 'Committee of Ministers on Energy Conservation' — shows that the junior minister concerned with conservation, Mr John Moore, is unhappy with Prime Ministerial attitudes. "Mr Moore" say the minutes "felt the committee could achieve little without the evidence of a political will from [Mrs Thatcher]".

The UK has not, in the past, come out well in international comparisons of measures taken to conserve energy, and it seems that the country is slipping further down the league. In its orgy of cutting spending, the government has:

- scrapped plans for 14 regional home energy saving advice centres (saving £500,000 a year)
- announced the closure in June of a scheme to help businesses improve inefficient boilers, insulate buildings, and install or improve combined heat and power schemes (saving £25 million over 2 years)
- halved the budget for assisting individuals to insulate their homes (saving £12.5 million).

But conservation needs more than the price mechanism alone. As the Secretary of State for Energy, Mr David Howell, said last year, there has to be "sensible and sensitive" government intervention or the energy conservation programme will be unbalanced. Pressure is

rising on Mrs Thatcher to take cognizance of this. Last month a syndicate of ten organisations including the Electricity Consumer's Council, the increasingly important Parliamentary Liaison Group for Alternative Energy Strategies, and other bodies with a long record of serious concern with the energy scene, requested that the Select Committee on Energy should turn its attention to the effect of spending cuts on the conservation programme; and it seems likely that it will follow their lead. The committee should do so, as a matter of priority, for the leverage on primary energy consumption of a pound spent on conservation is probably very great.

This is to say that a pound spent on conservation can save more energy than a pound spent on production can produce. According to some estimates, the factor can be as high as three. And judging by French plans (reported in *Le Monde*) to save 10 million tonnes of coal equivalent by spending £6 billion in industry, the French government itself — despite its ambitious nuclear programme — is aware of these relativities.

David Howell, opposition spokesman on energy, told *Nature* this week "Britain spends peanuts on conservation. We want a step-change in attitude". And the attitudes we need to change are those which see energy production as a matter of investment, but energy conservation as a matter of giving grants. The Central Electricity Generating Board, the Gas Board, the National Coal Board actually advertise to encourage us to use more energy: 'cook electric' advertisements appear regularly on television despite the fact that electric heating consumes three times the *primary* energy of the equivalent gas heater or cooker. Power stations are only one-third energy efficient.

As the Department of Energy memo has it: "If the funds spent on conservation replace the same or greater amounts spent on supply, should it matter that they are going to private companies or individuals rather than, say, the CEBB or the NCB?" Certainly not if the true blue objective of an energy policy is to ensure that the economy is unrestrained by constraints on energy supply. The economy can be released from energy costs more rapidly, more continuously (with less lumping of capital, unlike the £1 billion per nuclear reactor), and most likely, more cost-effectively. If you conserve energy, Mrs Thatcher, you conserve cash; and you conserve it rapidly, with payback times that are within the lifetime of your parliament. Is that not sufficient political incentive? □



The Antarctic

Managing krill in the Southern Ocean

THROUGHOUT next month, the thirteen signatories of the Antarctic Treaty will meet in private in Canberra, Australia to negotiate a convention to control the exploitation of krill — commonly regarded as the Antarctic's most immediately exploitable resource. Krill are small shrimp-like creatures which thrive in great abundance in the extraordinarily nutrient-rich cold waters of the Southern Ocean. They are central to maintaining the ecological balance of the Antarctic because they feed on abundant phytoplankton and in turn are fed on by many larger animals; in particular they form the staple food of whales. They are also increasingly being recognised as a potential source of protein for man. Hence the need for a convention which controls man's exploitation in such a way as to be non-discriminatory between different nations and at the same time maintain the local eco-system.

The Antarctic Treaty states have been negotiating the 'krill convention' — Convention on the Conservation of Antarctic Marine Living Resources in secret — for several years. Next month's meeting is hoped to be the last: agreement on the text of the convention should be reached, although ratification could take up to ten subsequent years.

The success of the convention, however, will also depend upon how acceptable it is to other nations not involved in its negotiation but nevertheless interested in reaping some of the benefit of the opening up of new fishing grounds. In particular, several developing countries have expressed fears that the convention only takes into account the interests of the

Antarctic Treaty signatories and other countries which are technologically capable of fishing in the extreme conditions of the Antarctic. The secrecy of the negotiations has not helped to dispel such fears and some observers, in particular the authors of a report published by the International Institute for Environment and Development (IIED)* last week, believe that agreement on the convention may spark off a wider international debate on the rights of access to the fishing grounds.

Reaching some sort of agreement just among the Antarctic Treaty signatories, however, will not have been easy. The key problem has been how to cope with the territorial claims of some nations to large slices of the Antarctic continent together with 200 mile-wide coastal strips of sea. Those signatories who have not laid claim to land do not recognise the claims of the others. The problem has therefore been to work out a 'bifocal' approach which leaves open the question of sovereignty over coastal fishing grounds.

The political problems of negotiating the convention are certainly great — but there are also many scientific problems to be solved before a system of good management for the Southern Ocean can be established. As yet there are no reliable estimates of the amount of krill which could be caught annually on a sustainable basis. There is also insufficient information on the interaction of different species living in the area.

The scientific problems implicit in the krill convention are different from those associated with other fisheries agreements

because of the nature of the convention's prime objective — the conservation of marine living resources. This is a totally new approach to fisheries control according to Dr John Beddington of York University who helped the IIED prepare its report. Previous agreements have mainly been concerned with the optimization of harvesting by man.

Although Dr Beddington regards the spirit of the convention on conservation as worthy, he is critical of its scientific approach. In particular, it does not take into account the fact that the net greatest annual increment in a population can be achieved only by removing its predators. In the case of krill this would mean killing whales — a contradiction of the convention's aim of allowing depleted whale populations to recover.

The IIED report recommends alterations to the wording of the convention to eliminate this contradiction. It also suggests that the scientific committee which is to be set up under the convention to provide scientific information, should maintain close links with independent scientists. However, the BIOMASS (the Biological Investigation of Marine Antarctic Systems and Stocks) programme, managed under the aegis of the International Council of Scientific Unions, should remain independent from the scientific machinery of the convention, it says. It should supply independent scientific information, free from any political constraints. **Judy Redfearn**

**The Management of the Southern Ocean* Barbara Mitchell and Richard Sandbrook, IIED, 10, Percy Street, London W1P 0DR. £2.

Energy

France plans one third nuclear energy by 1990

A confident forecast by the French Ministry of Industry has predicted that by 1990 France will have cut its dependence on oil to one third of its primary energy demand. It will do this, says the ministry, mostly by a dramatic ten-fold increase in the use of nuclear power, but also through contributions from gas and renewable energies — particularly solar and geothermal. The renewable contribution must increase four-fold in the next ten years to meet the ministries' targets.

France already makes considerable use of hot springs for household heating. All renewable resources (other than hydro) together now contribute only 1.5% of France's annual 323 million tonnes of coal equivalent energy demand, but, says an official communiqué of the council of ministers "certain renewable resources will come to maturity" in the next decade.

These projections considerably exceed those, say, of the UK, where energy

minister David Howell told the Select Committee on Energy recently that "the contribution from renewable sources in the year 2000 might be up to 10 mtce. My department now considers that this is, if anything, an overestimate". 10 mtce would be 2.5% of total energy use in 2000.

UK nuclear projections are also much smaller, a doubling to 9% of demand by 1990, compared to France's 30% by the same date. France must build one reactor every two months to reach its target.

According to *Le Monde* (4 April) the government expects that a £6 billion investment in energy conservation in industry by 1990 will save 10 mtce per year.

Moreover to encourage those living near the sites of projected nuclear power stations to accept their lot, electricity supplies will be set 12 to 17% cheaper for those living "in the vicinity" of the power station.

However not all are happy with the

plans. *Le Monde* quotes M. Michel Rolant, national secretary of the giant trade union CFDT, as saying that while it is desirable to reduce dependence on oil "it is dangerous to make this objective dependent on the imposition of electrical energy on all domains other than transport". The plans imply that in 1990 42% of French primary energy will be converted to electricity, 73% of that being produced by nuclear power stations. At present about 30% of French primary demand goes to electricity generation.

Moreover, said Rolant, France is taking a formidable bet on the good performance of the 66 or so nuclear reactors it will need in operation by 1990.

The cost of this programme — £26 billion according to Rolant — "is so large that it will make it impossible to follow a serious programme of energy conservation, or of the development of renewable sources . . ."

Robert Walgate

BIOPEN, the European-based genetic engineering firm, has had 600 "very good quality" applications for 18 post-doctoral positions in its new Geneva laboratory, Biogen's President, Robert Crawthorne, said last week. The applications come from throughout the world, but it seems likely that one third of the posts will be filled by applicants from continental Europe, one third from the UK, and one third from the US.

"The numbers we are looking for won't drain the pool of available talent" said Crawthorne "but we're looking for the best". Other companies may have had greater difficulty in recruiting, but Biogen has had something special to offer, he said: a chance to work with a scientific board of ten top academics in molecular biology (seven in Europe and three in the US) and "strong and open" links with the universities.

Biotechnology

600 apply to Biogen

Biogen is also planning greater links with European companies, particularly in chemical, petrochemical, and process engineering, said Crawthorne. The first cooperative agreement is expected shortly. "We already have a strong European shareholding: 36%, with 41% in the US and 23% in Canada. But we are not looking for shareholders any more." Rather, Biogen wants to learn from relevant industries what might be the most profitable products to develop with genetic engineering. "We've got expertise in recombinant DNA" said Crawthorne. "The problem is knowing what line to work on."

Biogen has found it very profitable to work with the US pharmaceutical company, Schering Plough, to whom it has licensed interferon production based on the successful cloning and expression of the leucocyte interferon gene by one of Biogen's scientific board, Professor Charles Weissman of the University of Zurich. "We want to extend such links" said Crawthorne, in Europe and elsewhere.

On interferon, he said Biogen was nine months from producing enough for clinical trials. A number of laboratories may undertake pilot plant production, including Professor Brian Hartley's laboratory at Imperial College, London, but decisions on such matters would be undertaken in consultation with Schering Plough. However, there were still a lot of hurdles, particularly over yields and purification.

Robert Walgate

Soviet Union

Moscow seminar meets

EARLIER this week, several non-Soviet scientists gathered in Moscow with about 50 of their Soviet colleagues for the fourth annual "International Conference on Collective Phenomena". These conferences, which are held outside the usual rigid framework of scientific exchanges, grew out of the weekly "Sunday Seminars for Refusniks", which began in 1973 to provide some kind of intellectual life for Jewish scientists who, having applied for a visa for Israel, found themselves denied professional employment.

Under their restricted conditions of work, without access to libraries, laboratories, computers and the like, it is, inevitably, the mathematicians and theoretical physicists who have the best chance of producing meaningful research. The majority of scheduled papers, therefore, were of a mathematical nature. Indeed, the hosts of the conference, Viktor Brailovskii and his wife Irina, are themselves both mathematicians.

One of the participants, Aleksandr Ioffe, since losing his academic post in 1977 has published several papers in foreign journals, including the *Transactions of the American Mathematical Society* and the *Comptes Rendus* of the Académie Française. In February, 1980, his foreign colleagues held a two-day mathematical symposium in his honour at Imperial College, London, where all but one of the papers were developments of Ioffe's work.

Ioffe, unfortunately, is something of an exception among the refusniks, most of whom can have little contact with the scientific community abroad. The "Collective Phenomena" conferences (as opposed to the more domestic seminars) were founded

precisely to overcome this feeling of isolation. This year's seminar was sponsored by the UK Institute of Physics, the New York Academy of Sciences, the French Physical Society, the French Institute of Chemistry and the Norwegian Physical Society, inspired, one may assume, by feelings of scientific solidarity rather than the hope that any major new research would be presented there.

The same solidarity, since the conferences began, has encouraged several western scientists each time to make the journey to Moscow, and to present papers of their own. Under Soviet conditions,

organizing an "independent" scientific meeting is liable to run foul of the authorities. Although some intending participants were refused visas, those who actually reached Moscow did not meet with the particular brand of bureaucracy directed at Dr Anthony Kenny, Master of Baliol College, Oxford, last weekend. His lecture, at the underground "Patocka University" in Prague (which met in the flat of Dr Julius Tomin, the dissident philosopher) was interrupted by the security authorities, on the grounds that his visa had been given for tourism, not for addressing other scholars. □

Spaceflight record-holder in orbit again

VALERII Viktorovich Ryumin holds the current record for orbital flight (175 days). His return last week to Salyut-6, where he established this record last year, has caused certain speculation abroad that the Soviets might be preparing a manned mission to Mars.

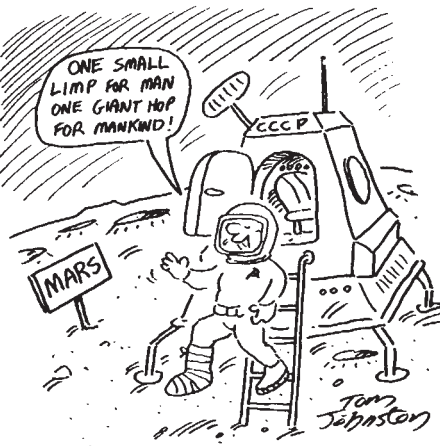
Such a mission, in fact, would conflict with all "classical" Soviet projects for space exploration. These, following Tsiolkovskii, the "father of Russian

cosmonautics" envisage the establishment of a permanent orbital station as a necessary preliminary to any deep-space mission. Such a station, it is understood, would be staffed by "several tens" of male and female crew and scientists for shifts of several months at a time. At least for this century, it is understood, lunar and planetary exploration will be left to the much cheaper automatic rover vehicles.

Nevertheless, Soviet space-medicine experts are showing considerable interest in the biological effects of long-term spaceflights. Last month, Academician Oleg Gazenko told TASS that the most important result to date was that no significant organic changes had been revealed which might limit the duration of future flights, and that people could work in space for six-month periods.

No doubt Ryumin, with last year's record-breaking flight to his credit, as well as a brief mission in October 1977 aboard Salyut-25, will prove an especially interesting "subject" — albeit a somewhat serendipitous one — during his current flight on Salyut-6. For, so far from being part of a Mars-preparation programme, he was, according to TASS, only included at a very late stage, when the original candidate, Flight Engineer Lebedev, injured his leg during final training.

Vera Rich



United Kingdom

18% photoefficiency claimed in BP energy research prize

FIXED photochemical energy amounting to 18% of visible light has been claimed by one of the recipients of a new energy research prize last week. The figure compares favourably with the electrical efficiency of solar cells, and is so far beyond the 3% normally regarded as the maximum achievable by plants that some specialists are highly sceptical about it.

The prize — £13,000 — comes from the oil multinational British Petroleum, which announced a £1.5 billion a year capital investment programme in its annual report published last week. "In the future" says the report "an increasing proportion of new investment, research and enterprise will be directed to activities additional to oil and gas". The BP energy prize is considered to be a way, said a BP spokesman, "to begin to get in on the act" on increasingly interesting energy research in the universities. But BP will not claim ownership of any of the results.

BP has awarded £39,000 to be divided between three groups from four UK universities. Similar awards will soon be made in ten other countries: Germany, France, Belgium, Denmark, Holland, Switzerland, Greece, Portugal, Canada and New Zealand. After a year's research, reports from the three groups in each country will be considered, and a national winner selected — who will receive a second year's grant and a cash prize of £5,000. These winners will also be considered by an international panel for an international energy prize of £10,000 to be awarded in July 1982.

The UK winners of the first leg are: Professor S J Pirt of Queen Elizabeth College, London, to develop an algal bioreactor for fixing solar energy; Drs H A O Hill of Oxford and I J Higgins of Kent Universities to improve the efficiencies of fuel cells powered by enzymatic reactions; and Dr R P Howson of Loughborough University to develop optical coatings for improved heat retention by windows. The prizes were awarded by a panel of five drawn from the Royal Society and the Fellowship of Engineering.

Professor Pirt hopes to build a 0.5m² collecting area 'unit' bioreactor, which would be deployed in multiples over a solar collecting field. The collecting surface will be tubular, and carry a continuous culture of *Chlorella* fed by ammonia, salts, and 100% CO₂ ("which" says Professor Pirt "we have discovered is not toxic, contrary to expectations").

The system has been developed in the laboratory. The next stage involves scale-up, and the application of microprocessor control to adjust nutrients and flow rates for varying sunlight intensity.

"Our target is to fix 14% of solar energy incident on the field", Professor Pirt said last week. "We have reached 18% photoefficiency in the laboratory" he claimed. Asked if this was not a very high figure, he said there had been no reliable and consistent data for photosynthetic efficiency. "Figures vary by a factor of four". (Pirt has published his results in *J. Chem. Tech. and Biotech.* 30, 25-34; 1980).

Photosynthetic processes were not at their most efficient at the CO₂ levels present in air, said Pirt. It had proved crucial to provide pure CO₂ to the organism "despite the fact that the books say more than 5% is toxic".

Pirt is "quite confident" that the capital costs of the reactor can be brought below those of competing solar technologies, such as solar cells.

One potential advantage of his closed system — where all external factors other than sunlight and temperature can be closely controlled — is that evolved oxygen can also be collected, and would become a by-product of the process. Moreover collected energy — in algal biomass — would be automatically stored, unlike the electrical energy from solar cells.

The biomass could be fermented to methane to provide natural gas; but fixed nitrogen would have to be recovered if the process were to be overall energy efficient. So "there are a number of other biotechnologies to be appended to this system", said Pirt, "before it is complete".

However, other biomass specialists are highly critical of Pirt's results. One pointed out that the currently accepted, "rock-hard" photosynthetic pathway proposed by Dr Robin Hill of Oxford 20 years ago required 8 photons to fix one CO₂ molecule. This leads to a theoretical maximum energy efficiency of 12%; in practice the record is sugar cane's 3%. But Pirt claims 18%, with a mixed culture of his alga and three heterotrophic bacteria.

"Pirt has claimed that one can produce 150 tonnes per hectare per year, but the best practice anywhere else is 30 to 50," said the biomass specialist. "He is doing biomass a fantastic disservice by making these claims" he said.

Pirt said last week that the accepted photosynthetic pathway "lacked any sound energetic data in support of it". His data "implies that we must look at what is wrong with the pathway" he said. "The gaps in our knowledge of it are really enormous. Otto Warburg [Nobel Laureate for medicine, 1931] always claimed that 4 photons per carbon was the correct figure. Others claim 12."

Robert Walgate

Hungary

Paks power station secured in concrete

HUNGARY'S nuclear power stations at Paks, the first generating set of which will go on stream in 1981, is so safe that not even an earthquake presents any serious danger. So said MP Miklos Vida, recommending a new nuclear energy bill to the Hungarian National Assembly last month. On behalf of the parliamentary committees which had dealt with safety questions, Vida noted that "the nuclear power station in the Armenian SSR which resembles the Paks power station easily withstood the force 5 earthquake . . . of 1976; according to estimates it can withstand force 9 earth movements".

Nor, said Vida, is there any fear of a nuclear explosion. "The system of construction of our nuclear power stations is such that it is a physical impossibility for a chain reaction to get, so to speak, out of hand". The only possible danger would be from "harmful radioactive materials finding their way into the environment". And "all necessary measures" have been taken to guard against this. The spent fuel "which represents the greatest source of radiation", will be stored "for some time", then sent back to the Soviet Union for reprocessing.

The remaining waste (liquids with a "high radiation content" and "discarded installations and parts") will be sent to the "isotope cemetery" established in a lenticular clay deposit near Puestoekszilagy, already used for radioactive wastes from industry and medicine.

For the Energy Ministry, Deputy Minister Gyula Szeker spoke of current fears regarding nuclear power, stressing that the western nuclear debate was encouraged by political, commercial and other interests which used "the revulsion felt for nuclear weapons" to "create a mood of opposition".

What Szeker did not mention was that the plans for the Paks power station depart from the standard Comecon doctrine that concrete containment vessels are unnecessary, and a capitalist ploy to raise costs. However "safe" the reactor, the Hungarians, it appears, will not neglect a little extra protection.

This somewhat more realistic approach to safety was also reflected in the speech of Imre Markoja, Minister of Justice. He admitted that even when all safety requirements are met "there can be some damage". Matters of responsibility and indemnity, he said, are therefore covered by the Bill. Although damage "due to exceptional events" in the course of operation of the reactor and the transport of nuclear materials is the least likely to occur, it is, however, the most dangerous. Indemnification in the case of damage, would be guaranteed by the state.

Vera Rich

NEWS IN BRIEF

French data systems sabotaged

PROGRAMMED tapes were stolen and fires were set at computer installations in Toulouse by two anti-computer groups last week. The urban guerilla group Direct Action claimed it had removed material "destined for use by the secret services" from the Toulouse offices of Philips Data Systems which it would later release. The second attack, a fire set at the rival computer company, CII-Honeywell-Bull, was claimed in a letter to *Liberation* by a group of "computer technicians well placed to appreciate the dangers to society," calling themselves the Committee to Liquidate or Neutralise Computers. The attacks followed a French government announcement of its plans to issue tenders for extra facilities to expand telematic memory banks for industrial, official and public use.

Jailed South African physicist in court

DR Renfrew Christie, the Oxford-trained nuclear physicist made his first court appearance last week to hear charges that he gave information about South Africa's nuclear programme to banned organisations. Christie, aged 30, has been jailed and held incommunicado without access to lawyers or friends under Section 16 of South Africa's Terrorism Act since his arrest shortly after he returned to take a job at the University of Cape Town last January. He is accused of intending to pass information about the location of South Africa's secret nuclear test sites, and a plan of the layout of the nuclear power station at Koeberg, to the African National Congress and the South African Christian Institute, two banned opposition groups. The trial will continue this week.

AGR plans approved

IN a reversal of her previous decision, Prime Minister Margaret Thatcher has been persuaded to approve plans for the construction of two British made advanced gas cooled reactors as part of the UK's £10 billion nuclear energy programme.

The Central Policy Review Staff and the Department of Industry argued strongly that the British nuclear industry could not develop competitively if the Prime Minister's preference for the American-built pressurised water reactor excluded the AGR. At stake are 3,000 AGR-dependent jobs in boilermaking, pipe and valve manufacturing, and engineering. Capital costs for the two proposed plants, at Heysham in Northwest England, and Torness in Scotland, have risen by 40% to \$2.8 billion in the past year, partly as a

result of increased public demands for safety. The current reactor design has several new safety features, including an enlarged reactor diameter for easier inspection and repair, and extra fuel channels to permit the reactor to maintain full power while operating at less corrosive lower temperatures. The plans also include a spherical concrete containment vessel instead of the cylindrical one.

Fewer Britons die at work

FEWER people were killed in accidents at work in the UK in 1978, according to the Health and Safety Executive Report* for 1978/79. The reduction in deaths — 498, compared with 648 in 1974 — is attributed to a greater concern about hazards at work and better prevention resulting from the 1974 Health and Safety at Work Act.

But fatalities among coal mining and railway workers show a disturbing upward trend. Mine deaths increased from 40 in 1977 to 63 in 1978; those on the railways stood at 48 for 1978, 14 higher than the average figures for recent years.

Approximately 20% of the HSE's total expenditure of £49 million was spent on research, testing and scientific support services for the year in question says the report. Extramural funding accounted for £2.6 million, of which two thirds were devoted to safety in the nuclear industry and to occupational health. Extensive research into the safety of pressurised water reactors and the hazards of exposure to plutonium were studied, safety procedures for genetic engineering were developed and the possible toxic or carcinogenic effects of a number of substances encountered in the workplace were investigated.

**Health and Safety Commission Report 1978/79. Available from HMSO £1.75.*

UMIST moves into biotechnology

THE University of Manchester Institute of Science and Technology (UMIST), in a move to become involved in biotechnology, has appointed its first Professor of Applied Molecular Biology. Dr Paul Broda, whose expertise is in the field of bacterial plasmids, will take up the post in the revamped Biochemistry Department of UMIST this September.

UMIST plans to back up the appointment of Dr Broda with a number of other new posts so as to assemble a team that will carry out fundamental research in areas of molecular biology that have potential industrial applications. One particular area of interest will be in antibiotic production from the streptomycetes. Another will be in the use of bacteria to derive end products of commercial value from lignin, a major waste product of the pulp industry.

Report says 55,000 UK smokers die each year

ON World Health Day last week a report issued under the auspices of the World Health Organisation warned of the risks of smoking, and singled out the UK as the worst sufferer. Some 55,000 there die prematurely each year as a result of smoking, says the report, compared with an over 18 annual death rate from all causes of 650,000. Some 41% of adult Britons are regular smokers; one in ten of these will die prematurely of causes directly attributable to smoking and one in four from diseases in which smoking is a major factor.

A study published last week by *Mintel*, a market research magazine, showed that while the number of cigarettes smoked in Britain fell slightly, from 125.9 billion, to 125.7 billion, from 1977 to 1978, the weight of cigarette tobacco smoked rose 10% from 197 million lb to 218 million lb. *Mintel* attributes this to the increasing proportion of sales taken by king-sized brands (10% in 1975, 50% in 1978 and a projected 70% in 1981). This rise in turn is attributed to the adoption in 1978 of EEC-wide tax rules, which favour larger cigarettes.

Academy recommends cut in formaldehyde exposure

CONSUMER exposure to formaldehyde gas, in particular that resulting from the use of urea formaldehyde (UF) foam for house insulation, should be reduced to the "lowest practical concentration" to avoid its irritant effects, according to a panel of the National Academy of Science.

In a report prepared for the US Consumer Protection Commission, the Academy's Committee on Toxicology refers to two European studies, one carried out in West Germany and the other in Denmark, as evidence that "there is no population threshold for the irritant effects of formaldehyde in humans."

The CPSC says that it has received more than 600 complaints from consumers living in homes insulated with UF foam. Reported symptoms range from eye and skin irritation to breathing difficulties, persistent nosebleeds, and nausea.

The Commission adds that the Academy report will help it to develop regulatory standards for consumer products containing formaldehyde. Both the commission and the formaldehyde industry are particularly concerned about the results of recent tests on laboratory rats carried out by the Chemical Industry Institute of Toxicology, which, although not yet completed, already show that 20 per cent of rats exposed to formaldehyde have developed an unusual form of nasal cancer.

FEATURES

Let there be light!

The recent growth of conservative religion in the US has injected new vigour into attacks on the teaching of evolution. **David Dickson** reports from Atlanta, Georgia

MORE than a hundred years after the publication of Charles Darwin's *Origin of Species*, it comes as a surprise to discover that almost half the adult population of the US believes itself to be directly descended from Adam and Eve.

Yet fifty years after a Tennessee courtroom witnessed the public ridiculing of divine creation, creationist beliefs remain as strongly held as ever.

The debating skills of Clarence Darrow in the famous Scopes 'monkey' trial of 1925 may have temporarily taken the wind out of the creationist sails. But today both supporters and critics agree that the creationist movement is growing in both strength and confidence buoyed by a rising tide of conservative ideology that is rapidly becoming a powerful force in American politics.

Many schools throughout the country, for example, are now required to teach creationist beliefs in parallel with evolutionary theory. And in at least six states — including Illinois, Florida, and New York State — legislatures are discussing bills which would make such practices compulsory.

Emotions on the issue run high. Most scientists continue to treat the creationist movement with derision and scorn. "These people are using glib salesmanship to sell an academic snake-oil for the general population" says Dr William Mayer, Director of the Biological Sciences Curriculum Study in Boulder, Colorado.

In the south and mid-west, however, the reaction is different. Here support for an absolute morality has strong appeal, basing itself on a literal interpretation of the Bible.

These regions have been fertile ground for creationist organisations which, in the words of Dr Henry Morris, director of the California-based Institute for Creation Research, seek to "reverse the dangerous drift of our country and its educational system into humanism, socialism, amorality and atheism".

Nowhere has the debate been fiercer than in Atlanta, Georgia. Last month the state legislature narrowly failed to approve a bill requiring that, wherever evolution is taught as part of a biology course in state schools, equal time should be given to creationist theories. Passed by the Senate,



God creates light: a creationist's view of evolution

the bill failed in the House of Representatives in the closing minutes of the last legislative session.

The state Department of Education had argued strongly against the bill, largely on the grounds that details of the curriculum should be left to local school boards. And among other critics the bill attracted the attention of the national 'atheist' organisation which had been responsible for having school prayers declared in violation of the Constitution.

The explicit involvement of the atheists has been a red rag to conservatives such as Judge Braswell Dean, one of the leading supporters of the creationist movement in Atlanta, who has accused the organisation of defending the 'monkey mythology of Darwin'.

To Judge Dean — and other creationists — evolution is an "animal fairy tale", based largely on a "superstitious" trust in chance in its belief that random mutations could have produced the current diversity of animal types. "Scientific creationism is far more scientific and less religious" says

the Judge, who blames the teaching of "humanistic" values in schools for social problems from rising crime rates to abortion, pornography and pollution.

Not all creationists are as strident. But they do share a common belief that the teaching of evolution implies a relativity in social values that undermines both traditional codes of morality in general, and the authority of the Bible in particular.

So far the creationists have had little success in convincing the courts of their case. In most instances the barrier has been the constitutional requirement that the state should not teach religious principles. These were the grounds, for example, on which the courts declared as unconstitutional an 'equal time' bill passed in Tennessee in 1973.

Undaunted by such setbacks, creationists are now trying to get round this problem in two ways. Firstly they argue that, since it is impossible to produce scientific 'proof' of evolution, it is no more than a hypothesis which might otherwise be described as a faith or religion.

Secondly, the creationists seek to convince the legislators that creationism can be legitimately called a scientific theory. And they distinguish what is now called 'scientific' creationism from both 'divine' creationism (responsible for the failure of the Tennessee bill) and even from 'biblical' creationism (ridiculed in the Scopes trial).

To defend this position, scientifically-trained creationists now scan the scientific literature for uncertainties, ambiguities and potential errors in the ideas of conventional evolutionists.

Gaps in fossil records and uncertainties in dating techniques, for example, are used to weaken confidence in what is called the 'evolution model'. And where science has disproved rigid immutability by experiment, creationists will now allow 'adaptive mutations' — but still deny that one species can evolve into another.

Most scientists reply that weaknesses in a rigid interpretation of evolutionary theory do not undermine their general confidence in the principles involved; and they accuse creationists of distorting the scientific method, and cause-effect mechanisms of scientific explanation, into a form which few laboratory workers would recognise.

Creationists see it differently. "Negative evidence against evolution is the same as positive evidence for creation", says Dr Morris.

Whatever the philosophical disagreements, few deny that creationism has found a ready audience outside the scientific community. In particular school curricula remain a major battleground for the creationists and their opponents, largely because these are expected to demonstrate compatibility with community values.

A typical case is Cobb County, a suburb of Georgia. Last December biology courses were removed from high school graduation requirements in an attempt to resolve a conflict originating 18 months earlier when the local school board passed a resolution requiring evolution and creation to be given 'balanced treatment' in school courses.

Heated objections from science teachers and some community leaders led to eventual agreement that a voluntary course on 'comparative theories or origins' would be offered for high school students and that Darwinian evolution would not be included in any school biology course.

The Cobb County dispute is typical of conflicts now being fought out across the country. Most have in common the fact that, as Dr Dorothy Nelkin of Cornell University points out in her book *The Science Text-book Controversy*, the loudest critics of evolution are not from lower-class, uneducated backgrounds, but tend to be middle-class, technically-trained citizens.

Cobb County, an area of rapid post-war growth centred on the communications and aerospace industries, typifies what

Nelkin refers to as the 'paradox' that fundamentalist beliefs tend to flourish in those parts of the country which have recently become centres of high technology industries (for example Southern California and Texas).

Also typical of the Cobb County controversy is that it has taken place against sharp religious differences in the local community. Such differences have traditionally polarised around conflicting ways of interpreting the Bible. Indeed rather than a conflict between religion and science, the current disputes centre around the legitimacy of different forms of knowledge, whether used to support religious or scientific ideas.

In Atlanta, strong support for creationism has come from the more militantly conservative churches belonging to the Southern Baptist Convention, members of the currently dominant group within which is now the largest protestant sect in the US. In contrast, more moderate theologians are among the most persistent critics of biblical literalism, and have been

'Evolution is an animal fairy tale' — Judge Dean

among the most keen to keep creationism out of schools.

Further controversy has focused on the role of the Institute for Creation Research, (ICR) and was ignited when it was discovered that the 1978 resolution passed by the school board was almost identically worded to a model resolution circulated by the institute.

ICR is a division of the Christian Heritage College, a group which split from the more liberal American Scientific Affiliation in the 1950s largely over conflicting views on whether the Bible should be taken literally or metaphorically.

Since then, the institute has become a powerful centre for the teaching and propagation of creationist ideas which like other modern evangelical organisations, it has done with spirited fervour.

Institute members were also influential in setting up in 1963 the Creation Research Society, an organisation consisting of over 600 individuals with postgraduate degrees in scientific or technical subjects concerned with developing a 'scientific' critique of evolutionary theory — and arguing the case for a 'scientific' alternative.

Although denying any orchestration of local creationist movements, Dr Morris admits that an important role of ICR is to provide model resolutions and educational literature to groups prepared to support the contention that "teaching evolutionary theory alone is contrary to academic freedom, civil rights, and the freedom of religion".

Ironically much of the creationists' critique of the values implicit in evolutionary theory are similar to those coming from a very different perspective

which identify Darwin's ideas with the political norms of Victorian capitalism. But despite apparent similarities, the antagonisms are sharp.

Professor Richard Lewontin of Harvard University, for example, describes the recent rise of creationism as reflecting a "wave of anti-scientism", suggesting that its rejuvenation is "because intellectuals are identified as the allies of social movements" that challenge traditional power relations in society.

Conversely Dr Morris points out that many of the more politically radical evolutionists are among those questioning the more dogmatic aspects of Darwinian theory. But he castigates them for doing so "with prejudices tied to Karl Marx rather than Adam Smith".

It is difficult to judge the impact of the creationist movement. Some argue that, by causing publishers and teachers to drag their feet over introducing evolution into the class-room, creationists — in the words of one educator — "have had a tremendous impact on students in denying them access to scientific knowledge".

Others, however, are more prepared to be flexible. They use arguments from the sociology of knowledge to defend the interpretation of science as a belief system; and side with the creationists in their criticism of the more dogmatic assertions of some scientists and textbook writers.

Of greater concern to educationists are the implications of a movement which appears to attack abortion and homosexuality and defend the 'free world' with the same degree of dogmatism that it criticises in others.

Jeremy Rifkin, author of *The Emerging Order*, warns that what is now a religious movement, reacting in part to the 'idolatry' of scientific and technological truths, could turn into the opposite. "Christian doctrine, made an adjunct to right-wing and capitalist policies, could provide the necessary self-imposed order that a fascist movement in America would require to maintain control over the country during a period of long-range economic decline" he writes.

Creationists are used to responding in kind. Dr Morris argues that both fascism and communism have their roots in evolutionary thinking. He attacks the "humanistic and socialistic" biases of textbook publishers who refuse to mention scientific creationism. And he claims that creationist ideas have unfairly come under "bigoted pressure from the liberal news media".

Whatever the words used, the bitterness of the Georgia conflicts indicate that creationist ideology is unlikely to disappear. And it seems only a matter of time before some state requires creationism to be taught wherever Darwinism is invoked and another court is asked to rule on the respective definitions of science and religion. As Dr Mayer says, "It is all part and parcel of the signs of the times". □

Patenting nature's secrets and protecting microbiologists' interests

THE patenting of inventions in microbiology, is arousing unusual interest the United States and Europe. **Stephen Crespi**, Patents Controller at the UK National Research Development Corporation examines recent developments in Europe

THE US Supreme Court is hearing the appeal brought by Dr A Chakrabarty in 1972 over the rejection by the Patent Office of an attempt by GEC to patent an oil-consuming strain of *Pseudomonas* obtained through genetic manipulation. And in Europe, the European Patent Office has recently changed the mechanism, called Rule 28, for ensuring public availability to third parties of new strains of microorganism deposited in culture collections for patent purposes.

The legal controversy aroused in the US by the long-running Chakrabarty case has not fully erupted in Europe because there has been no comparable test case. But Europe must view developments in the US with interest because in a world in which legal systems borrow from one another the outcome in the US may affect that in Europe.

Some of the older American decisions show judicial condemnation of attempts to patent nature's secrets where living organisms are involved — although patents were allowed for meritorious discoveries of inanimate products of nature, vitamin B₁₂ being one of the most celebrated examples.

In Chakrabarty the fundamental point is simple; is a living organism which otherwise complies with legal requirements for patentability nevertheless disqualified because it is alive? In answering this question the court has to consider how to accommodate the product of genetic engineering in a patent system based on the models of classical physics and chemistry. However, if the decision of the Supreme Court later this year is limited to man-made organisms which are the result of genetic engineering it will be disappointing after so much effort has gone into dealing with the broader issue.

To put these developments into perspective it is worthwhile summarising the possible categories of patentable invention in microbiology. Microbiological processes have long been recognised by courts as suitable for process patents; and the newer patent statutes in Europe specifically refer to them. Patent claims may be presented for inventive developments of any of the methods of microbiology which serve a useful economic purpose such as improvements in culture media, culture conditions and the choice of strain used. Processes in which the sole novelty lies in the use of particular strains,

especially newly developed strains, have become conventional subjects of "process-of-use" patents.

Products produced by microorganisms such as antibiotics and enzymes have for a long time been patentable in the form of product claims where the products were novel and where the patent law of the country concerned permitted the so-called product-*per-se* claim, i.e. not limited in scope to a particular process. Where the product was not new, and novelty resided only in the process of making it, it was customary to use both process claims and also, where possible, the product-by-process claim, i.e. a claim to the product when made by the particular process. However, it is with the other type of possible product claim, the claim to the new strain of microbial cells themselves, that tension has arisen between applicants and examiners. Claiming the microbial biomass as a useful end product was generally acceptable but claims directed to the new strain *per se* seem more recently to have raised the kind of philosophical objections in the Chakrabarty case.

The traditional attitude of the UK patent system towards patent protection involving the use of living matter has been cautious but attentive to the needs of industry and applicants for patents. The following principles have operated: the concept of "manufacture" as essential for patentability; the fact that many living substances can be bought and sold like other commodities; giving the benefit of the doubt to an applicant where the law was uncertain but where the decision could be tested by a higher authority if any contestant so wished. The test of "manufacture" was whether the product was itself a manufacture or could be applied to manufacture, or whether the process was technical in nature as distinct from the establishment of conditions under which the organism was developed by essentially biological laws. Long before the subject attracted public comment, yeast manufacturers took out patents for new yeast strains, presumably because the alternative of trade secrecy was non-existent where the live microorganism was itself the item of commerce. British patents were also granted over the past twenty years for other new strains, cell lines, and attenuated viruses intended for vaccines. The *Fusarium graminearum* strains

intended as sources of single cell protein were patented without challenge in the UK in 1974 although held unpatentable in Eire and Australia by official ruling a few years later.

The British Patent Act of 1977 was brought in line with the European patent law of 1978 and practice should run in parallel under both systems. Recent informal discussions with EPO officials suggest that product-by-process claims to microorganisms will be accepted and that the decision on the patentability of unrestricted product claims to the strain *per se* will be taken soon. Much may depend on the circumstances of each case.

The number of patent applications on genetic engineering procedures reaching the publication stage is increasing. Once Patent Office Examiners have fathomed the extraordinary complexity of the subject there should be a spate of patents granted for techniques. It is difficult to see why recombinant DNA plasmids should be treated any differently from other chemical substances but whether claims to plasmids and transformed strains are obtained remains to be seen. What commercial value do these claims have and can they be policed inside the competitor's factory gates? We cannot yet judge the value of these to the innovators especially with so little experience of their usefulness. Therefore instead of being negative or restrictive we should explore ways in which the patent law can encourage these new areas of research.

What is the objection to patenting living matter? Some condemn it by asking where it will end and they argue that a logical extension to higher life forms supports their view. But law and logic are not identical and this argument has to overcome two objections.

The first is that one class of higher life forms is already protected, namely plants by the plant patents in the US and plant breeders' rights elsewhere. The latter are distinct from patent protection and show an interesting difference: the plant breeder can control not only the commercial marketing of the reproductive material of the new variety but also its subsequent multiplication whereas normally once a patentee has sold the product covered by the patent he cannot control it further. Secondly, patent legislation in some countries (including that of the European Patent Convention) specifically excludes patents for plant and animal varieties.

The patent law exists to benefit research and its financial supporters. If it becomes necessary to draw an arbitrary line between what living organisms are or are not patentable it is within the wisdom of judges

and administrators of the law to do so in a way which supports technology and is generous to the inventor without harm to anyone else. Borderline cases will arise where the meaning of the terms 'plant' and 'animal' will have to be looked at closely but the constraints against patents for higher life forms are clearly built in to the present laws, at least according to the European model.

In presenting their survey of the patent law to the Supreme Court, the US judges of the lower court have approached the question on the pragmatic ground of usefulness to industry where they see no distinction between living microorganisms and chemical elements and compounds. If it is socially acceptable and desirable for pharmaceutical companies to develop new microorganisms and produce products containing living material such as 'live' virus vaccines there can be no reason for restricting the patent cover available for these innovations.

To erect obstacles to patent protection is to encourage the secrecy which the patent law is designed to discourage. The patent system has in recent years become much more open especially with its emphasis on early compulsory publication of patent applications which in the past often remained confidential in the Patent Office for a long time. However, the emphasis on early publication has caused problems for microbiologists because a new strain of a microorganism must be available to third parties at the same time as publication. The European Patent Convention of 1973 set the trend on this point.

One of its regulations, Rule 28, had said that a new strain must be deposited in a culture collection before a European patent application could be filed properly and also insisted on the accessibility of the strain to others, subject to a few conditions, on publication of the application 18 months or so after the priority date. This contrasted with US and Japanese patent law where release of the strain is obligatory only when an enforceable right is obtained.

However, the European Patent Office has recently modified Rule 28. Availability of the strain to third parties can now be restricted between first publication of the application and the eventual grant of patent rights. During this time the applicant will be able to limit access to the strain to an independent expert acting on behalf of third parties but bound by certain conditions including that of not passing the strain out of his hands.

This improvement of the rule concludes over six years of effort by European industry and others to persuade the authorities that unrestricted availability of the culture before any rights are granted involves loss of control at too early a stage. This has been one of the first controversial questions tackled and solved by the EPO since it began operation in June 1978. The decision to change the rule has anticipated

Evolving ideas

EVOLUTION is not really in trouble, of course, it has never been healthier. It has gone into the computers. This is a sad business when one thinks of the halcyon days about a century ago. It is true that Darwin, the brooding sage, was a recluse at Down, but his supporters were having a wonderful time. We have been reminded of this in the splendid television series 'The Voyage of the Beagle', and especially by the exciting re-enactment, on this programme, of the confrontation between Bishop Wilberforce and Thomas Huxley. The question asked by the Bishop was on which side Huxley claimed descent from the apes. Today we might reply that the maternal line of inheritance has a slight edge because mitochondria probably travel with ova. However, Huxley's thunderous response was directed personally at Wilberforce, whereupon it is said, a young woman fainted. Ironically, Queen Victoria, the supreme head of Wilberforce's church, carried a mutant gene for one of the blood-clotting globulins (Factor VIII), for the male haemophiliacs among her descendants provided a tragic and classic example of the ruthless effects of natural selection.

Evolution then set forth for many years on an adventure among fossils of extinct animals and plants. Pterodactyl, *Tyrannosaurus* and *Archaeopteryx* became household words. Descriptions of the bones of our ancestors were often in the news. The Scopes trial put evolution into the field of entertainment. Next, biochemists devised methods for determining the sequences of amino acids in proteins. It became possible to measure evolutionary divergence numerically in terms of amino acid differences between similar proteins in various species. Haemoglobins of gorillas, chimpanzees and human beings were distressingly similar, but widening differences were found in other species. The fun was disappearing from evolution, but worse was to come for the classical taxonomists. Incredibly rapid new methods were perfected for measuring long sequences of nucleotides in DNA, and the results, photoreduced to near-illegibility, appear in *Nature* almost every week.

A few years ago, we could examine

the UK report on biotechnology published last week (see *Nature*, 10 April, page 502) which strongly criticised the lack of protection under Rule 28.

Rule 28 was also updated to conform to the corresponding rule in the Budapest Convention of 1977 which provides that deposition of the strain in a single officially recognised culture collection will suffice for the individual, national procedures.

The detailed application of the independent expert idea remains to be



THOMAS H. JUKES

phenotypes but we never expected to be able to read genes. Now every sequencer can become a computer-aided evolutionist. Viruses evolve just like entire organisms; in fact, genes in simian virus 40 and the polyoma virus have diverged even further from each other than genes for the alpha and beta chains of haemoglobin, which have spread far apart in the 500 million years since they separated from a common ancestral molecule. But who knows how fast viruses evolve? They leave no fossilized imprints in the rocks as guideposts of their age.

To look for a gene in DNA, you scan for 'open reading frames'. These are regions in nucleotide sequences that are free from occurrences of TAA, TAG and TGA: the 'stop signals' in protein synthesis. It is becoming quicker to find new proteins in DNA sequences than to separate them from protoplasm. It is even possible to find genes that are no longer in use. Phil Leder recently called such a gene (for a mouse alpha haemoglobin chain) a 'rusting hulk', because it had accumulated so many changes, including deletions and insertions.

All evolutionary changes result from inherited changes in DNA molecules. It is difficult to get emotional about alterations in the linear arrangement of A, C, G and T. The spiritual descendants of Bishop Wilberforce must find it rather dull to argue with computer programs.

worked out but the intention is that the expert will carry out experiments on behalf of third parties and potential opposers of the patent to test the patent disclosure and make an evaluation of the invention. The expert will be chosen by agreement between the applicant and the party requesting the strain or from an official list of recognised experts. Consequently the services of eminent microbiologists will be in demand and experts who might be willing to undertake this role are being canvassed. □

CORRESPONDENCE

Nuclear technology is not the threat

SIR,—Mr A. B. Lovins' article (28 February, page 817) confines itself almost entirely to a discussion of the use of reactor grade plutonium in the construction of bombs, either by terrorist groups, or by governments. I do not believe myself to be enough of an expert on these esoteric branches of industrial chemistry and nuclear physics to comment usefully on the technical aspects of the article, and in any case I would prefer to concentrate on some implicit non-technicalities.

First, there is the claim, not made here by Mr Lovins, that the increased availability of dangerous materials will increase the risk of their misuse. While there is some truth in this depressing doctrine, I feel that the size of such risks is dominated by the proportion of people who are keen to use terrorist methods, rather than by the technology of their particular community.

No, the question surely is whether we are to give up clean, reliable, safe, economical and enormously useful techniques simply for fear of what the malevolent may do. If we are, where does this line of argument stop? Do the Friends of the Earth want us to give up aviation, liquid natural gas, or genetic engineering? And if not, why not? Or, to put it another way, should Adam have stopped Eve from eating that Fruit?

In addition, I have some specific quarrels with Mr Lovins' article. Who, for instance, were the "three high-technology nations" whose ministers were so misled? Again, no-one associated with the anti-nuclear movement is entitled to complain about "lost, oversimplified or garbled" advice. And what on Earth does he mean by saying that "the implication that the effects of even a crude, minimal 0.1-1 kton explosion would be tolerable for a free society is at best disingenuous"? Non-free societies are entitled to proceed, apparently.

Without making any statement about the possible carnage in such an explosion, I would like to point out that conventional accidents sometimes take thousands of lives in free societies, and presumably in non-free ones too. The multiple standards of response to such disasters is sadly but glaringly highlighted by the failure of the Morbi dam in Gujrat State, India, in August. It seems to have taken some hundreds of lives at least, and passed almost unnoticed in the Year of Three Mile Island.

Yours faithfully,

J. F. CRAWFORD

Klein Doettingen,
Switzerland.

Comparing the diets of laboratory animals

SIR,—It has been demonstrated several times that toxicological experiments performed in different laboratories are liable to unexplained variations in result. Amongst the possible reasons for this observation, are differences in diet of the experimental animals (see for example Hathcock and Coon *Nutrition and Drug Interrelations*, Academic Press, New York 1978). Of 89 randomly-chosen recent publications in the fields of oncology, pharmacology and toxicology, more than half failed to specify the diet fed to their laboratory rodents. Often considerable portions of the 'methods' section were description of the animals, sources of

chemicals used, etc, but the nearest any of the authors got to specifying their diets was to give the manufacturer's name. This implies that the authors consider stock diets to be standard. However, my survey of 24 UK stock diets for mice and rats has shown that for each nutrient, very large differences in the diet composition exist. Amongst nine minerals, the coefficient of variation ranged from 16% to 80%, with a median value of 28%. Similarly for 12 vitamins, the range was from 20% to 91%, with a median of 62%.

The survey was based on the best information available, but in all cases it is a calculated figure only, based on the ingredients of the diets. Only one British manufacturer has been prepared to supply details for one diet of average analyses after production. Some manufacturers at extra cost will provide analytical details to customers on a batch basis, and it is these figures that are required for a scientific evaluation of diets on a comparative basis. The majority of nutritionists use semi-synthetic diets, rather than stock diets, and it is perhaps for this reason that their attention has not been focused on diets used by non-nutritionists and why so little nutritional work has been done in this field.

It is the intention of the Laboratory Animals Centre to try to obtain information about actual analyses, for nutrients and contaminants, on stock diets worldwide. It would then be possible to help those who are trying to explain differences in results between similar experiments in different laboratories, in which it is suspected that dietary variation might have been important. In the furtherance of this aim, it is suggested that diet manufacturers and scientists who have relevant analytical data should contact the Laboratory Animals Centre to discuss the possibility of collaboration in such a study.

Yours faithfully,

A. WISE

Medical Research Council
Laboratory Animals Centre
Carshalton, Surrey, UK

Establishing the case for the Maunder Minimum

SIR,—The News and Views contribution of 31 January (page 427) entitled "Was there a Maunder Minimum?" projects a superficial understanding of the meticulous work by solar astronomers, historians, climatologists, palaeobiologists, and scientists from a host of other research disciplines supporting a marked diminution in solar activity during the 17th century.

Table 1 of the article shows that there were nine naked eye sunspot records during the decade beginning 1610, six in the 1620s, nine in 1639, two in the 1640s, three in the 1650s, one in the subsequent decade, then none, and finally one in the 1680s. Taken at face value, the new *fang chih* records would seem to lend support to a dramatic reduction from about 1640 in the number of sunspot groups or individual spots large enough to be seen by the naked eye under suitable atmospheric conditions. But a direct interpretation of such limited data is superficial. As your article stresses, the *fang chih* do not present an unbiased sample of data, any more so than the records in the official dynastic histories. At least for the latter one has some check on completeness of the data by looking at the continuity of official records of other astronomical phenomena. However, because

of political, social and astrological influences, oriental sunspot records are by themselves of little value; as previous investigators have been careful to stress, they provide at best merely circumstantial evidence for excursions in solar activity. There is no point arguing for or against a Maunder Minimum on the appearance or absence of naked-eye sunspot records from the orient when post-telescopic (1610) occidental records, free from political-astrological influences, are available and when a wealth of other indicators of solar variability have been investigated. Yet none of these are mentioned, the reality of the Maunder Minimum being called into question on the basis of a few provincial oriental records of doubtful reliability.

The leading late 17th century observational astronomers lamented the scarcity of sunspot activity. Here we have more than just a chance mention of detections, but professional observers insisting that it had been a decade or more since any spots had been seen. Spots there certainly were during the Maunder Minimum (1645-1715) but greatly reduced in frequency and extent. Thus, for example, Stephen Gray noted in 1705 "the Sun was much more productive of (spots) in (Galileo's) time than it has yet been in ours, and the Regions Producing them had a far Greater extent". However, if one was hesitant to accept the direct evidence of contemporary scientific writings, indirect indicators provide near-conclusive proof of reduced solar activity. A history of the flux of cosmic rays, related to solar activity (and thence to sunspot occurrence) can be constructed from isotopic anomalies preserved in tree rings and ice cores. All such investigations reinforce the case for the Maunder Minimum.

Yours faithfully,

DAVID CLARK

Royal Greenwich Observatory, Hailsham,
Sussex, UK

F. R. STEPHENSON

University of Liverpool

For any readers wishing to follow up the original Chinese language paper on which the *News and Views* article was based the reference is *Nanjing Daxue Xuebao (Ziran Kexue Ban No. 2/31/1979 and not 1976 as printed: ED*

Lead petrol additives

SIR,—In an otherwise valuable contribution to the debate on the health impacts of childhood lead exposure, M.R. Moore (24 January, page 334) may have unnecessarily compromised the case against leaded petrol by conceding significant economic benefits.

It has long been accepted that lead petrol additives are a major contributor to combustion chamber deposits responsible for increased octane rating demand. Measurements of this effect (e.g. G. Cornetti *et al.*, *The Journal of Automotive Engineering*, June 1971, 8-14) have demonstrated fuel efficiency penalties that are greater than the gains made possible by increased compression ratios with high octane, leaded petrol.

Yours faithfully,

DAVID R. L. DAVIES

Centre for Resource and Environmental
Studies,
Australian National University,
Canberra

NEWS AND VIEWS

Interferons and the immune system

from Barry R. Bloom

The current exuberance about interferon reflected in these pages referring to interferon as "potentially the wonder drug of the 80's" recalls the comment of Michael Faraday on being questioned sceptically by Gladstone on the uses of his research on electromagnetism, "One day, it will give rise to a great industry upon which you can levy taxes." Why such enthusiasm? At least part of the answer lies in the recent appreciation of the regulatory effects that interferons exert on the immune system.

Interferon has been defined by its antiviral effect and has been considered almost exclusively as a potential antiviral agent. (The work of Gresser on the other biological effects of interferons stands as an heroic exception (*Cell. Immun.* **34**, 406; 1977).) Yet evidence for its efficacy in acute virus infections has been relatively disappointing until recently although it does seem a promising treatment in hepatitis, rabies, varicella zoster, herpes keratitis and rhinovirus (common cold) infections. But the most significant boost to the budding interferon industry comes from trials of interferon in cancer patients, initially stimulated by the studies of Strander and his colleagues on osteosarcomas, and reinforced by preliminary findings in non-Hodgkins lymphomas, breast cancer and multiple myeloma. So far, however, the optimism is based on published results in fewer than 50 patients in trials that have not been randomized, blinded or well controlled. As more and cheaper interferon becomes available, larger trials will be carried out.

The principal approach to studying both

the antiviral and antitumour effects of interferon has been to look at the direct effects of interferon on appropriate tumours or virus-infected target cells. Various metabolic pathways activated by interferon have been elucidated, involving 2' 5'-polyadenylate synthetase, protein kinase and nucleases which may play a part in blocking initiation of viral or host protein synthesis and in breakdown of viral and host mRNAs (see *Nature* **282**, 364; 1979). One probable consequence of these regulatory changes is a relatively general inhibition of replication of various normal and neoplastic cells. *In vivo* experiments confirm that interferons inhibit tumour growth in several animal models, but they can also suppress liver regeneration in partially hepatectomized mice and even inhibit growth and development of suckling mice.

In the light of the general growth inhibitory effects of interferons, it is not surprising that the first effects of interferon on the immune system to be studied were its immunosuppressive activities. Several laboratories showed that interferons can inhibit antibody formation to thymus-dependent (sheep erythrocytes) and thymus-independent antigens *in vivo* and *in vitro*. In this regard type II 'immune' interferon (produced by T cells) has been found to be 10-100 times more suppressive than type I interferon (the type produced by leukocytes and fibroblasts) (Sonnenfeld *et al. Cell. Immun.* **34**, 193; 1977). Interferons suppress mitogenic responses of both T cells and B cells *in vitro*. Indeed, interferons seem to be responsible for mediating the suppression of the concanavalin A-activated human T suppressor cell (Kadish *et al. J. exp. Med.* in the press). In most studies *in vivo*,

relatively large amounts of crude interferons have been used, and while suppressive effects were observed, interferon can hardly be regarded as one of the more potent immunosuppressive agents available. Indeed, little evidence for suppression, of antibody formation at least, has been seen in patients receiving interferon therapy for cancer or virus infection. Although interferon suppresses graft-versus-host reactions in animals, its usefulness in human bone marrow grafts is limited because its growth-inhibitory activity has blocked bone marrow proliferation and reconstitution. Thus the balance between useful immunosuppression, and unacceptable side effects may be delicate.

Interferon and NK cells

In 1975 it was first recognized that normal spleen cells of some strains of mice were cytotoxic *in vitro* selectively for neoplastically transformed cells (Kiessling *et al. Eur. J. Immunol.* **5**, 112; 1975). The cells responsible for this spontaneous cytotoxicity were termed 'natural killer' (NK) cells and are quite distinct from conventional B cells, T cells or macrophages (*Transpl. Rev.* **44**, 1979; *Natural Cell-Mediated Immunity against Tumours* (ed. Herberman) Academic Press, in the press). Athymic nude mice spleens possess increased NK activity, whereas cells from individuals with Chediak-Higachi syndrome or its murine analogue, *beige* mice, are almost totally deficient. The provenance and fate of NK cells remain unknown. NK cells bear surface markers characteristic of the T-cell lineage (Minato *et al. Ann. N. Y. Acad. Sci.* in the press), but also seem to share characteristics in common with

Barry R. Bloom is Professor and Chairman in the Department of Microbiology and Immunology at the Albert Einstein College of Medicine of Yeshiva University, New York.

macrophages (Lohmann-Matthes *et al. J. Immun.* **123**, 1883; 1979). Alternatively, they may represent an independent lymphocyte pathway. It may well be that NK activity is mediated by more than one cell type in different systems.

Interferon seems to be the key immunoregulatory signal for NK activity. Several independent studies earlier suggested that inducers of interferon *in vivo* such as poly(I):poly(C) or tilerone augmented NK activity in spleen cells (Gidlund *et al. Nature* **273**, 759; 1978; Djeu *et al. J. Immun.* **122**, 175; 1979). Surprisingly, known immunological adjuvants such as BCG or *Corynebacterium parvum*, which in some models have been shown to enhance resistance to tumours in both conventional and athymic mice also augment NK reactivity (Wolfe *et al. Nature* **252**, 584; 1976). It has now formally been established that interferon, including chemically-purified material, can markedly augment cytotoxicity of NK cells for appropriate target cells (Trinchieri *et al. J. exp. Med.* **147**, 1314; 1978; Zarling *et al. J. Immun.* **123**, 63; 1979; Minato *et al. op. cit.*).

Lymphoid cells are not only regulated by interferons but actively produce them. Type II or 'immune' interferon is produced by mitogen or antigen-stimulated T cells (Epstein in *Interferons and their Actions* (ed. Stewart) 91, CRC Press, Cleveland, 1978). Interaction of a non-B, non-T cell population of human peripheral blood cells with a majority of human tumour cell lines tested *in vitro* resulted in the production of type I leukocyte interferon, and there seemed to be a general correlation between the susceptibility of tumour cells to lysis by NK cells and the tumour cells' ability to induce interferon production by NK cells (Trinchieri *et al. J. exp. Med.* **147**, 1299; 1978). Several laboratories have shown that NK cells bind to appropriate target cells, and have used specific adsorption to targets to enrich for human NK cells and examine their morphology and characteristics (Timonen *et al. Cell. Immun.* **48**, 133; 1979; Roder *et al. J. Immun.* **121**, 2509; 1978; Jensen & Koren *J. Immun.* **123**, 1127; 1979). After the binding of a relatively homogeneous population of medium-large lymphocytes with striking cytoplasmic granules to susceptible target cells, these bound presumed NK cells rapidly become immunofluorescence-positive for interferon (Saksela *et al. Annl. N. Y. Acad. Sci.* in the press). While it has not formally been established that the interferon-producing cells are the cytolytic ones, the correlation between binding, interferon-positive immunofluorescence and cytotoxicity for the targets is striking.

Interferon could augment NK activity in two ways. (1) It might activate existing NK cells, like the activation of macrophages by lymphokines, or (2) it could trigger the differentiation of a precursor to the NK cell into an effector cell. It has recently been

shown in nude mice that both NK cells and interferon-producing cells are susceptible to anti-Ly 5 antibodies (Cantor *et al. Immun. Rev.* **44**, 3; 1979) and anti-Qa 5 antibodies (Chun *et al. J. exp. Med.* **149**, 426; 1979). When NK activity was depleted by treatment with anti-Ly 5 + complement, interferon restored it within 3 h, and the induced effector cells expressed the Ly 5 surface marker (Minato *et al. op. cit.* In contrast, treatment with anti-Qa 5 + complement eliminated NK activity which could not be restored by addition of interferon, and hence serves as a marker for the NK precursor. Precisely analogous results were found using human cells depleted of functional NK activity by selective adsorption to target cells which could be induced to become cytolytic after interferon treatment (Saksela *et al. Scand. J. Immun.* **10**, 257; 1979). Thus activation of NK cells by virus or appropriate tumour cells initiates a positive feedback circuit mediated by interferon resulting in rapid differentiation and amplification of NK function.

Role of NK cells in resistance to tumours

What is the evidence that interferon-induced NK cells have any significant immunological role *in vivo*? The mechanism by which interferon exerts antiviral effects *in vivo* should be re-examined following recent experiments which show that although human interferon did not protect monkey cells against vaccinia virus infection *in vitro*, treatment of monkeys themselves with interferon effectively prevented the growth of locally inoculated vaccinia virus (Schellekens *et al. Nature* **278**, 742; 1979). These results strongly indicate that in this, and perhaps other systems, interferon may exert its protective effects not by acting directly on target cells but perhaps through its effects on the immune system. There is evidence for a significant role of interferon in natural resistance to acute infections (Gresser *et al. J. exp. Med.* **144**, 1305; 1976). Surprisingly, the interferon-NK cell system may be far more important in resistance to persistent virus infections and tumour cells.

As few as 10–100 HeLa or BHK cells produce tumours in all athymic *nude* mice; yet 10⁶–10⁷ of the same cells persistently infected with any of six RNA viruses fail to grow as tumours (Minato *et al. J. exp. Med.* **149**, 1117; 1979). The persistently-infected tumour cells were killed by NK cells *in vitro*, whereas the uninfected parental tumour cells were not. One variant of BHK cells persistently infected with VSV was an excellent inducer of interferon and capable *in vivo* of augmenting cytotoxicity to irrelevant target cells, but was itself resistant to NK killing *in vitro*. This variant formed disseminated tumours and metastases in *nude* mice.

Important evidence for a role for NK cells in restriction of metastases has

recently been provided by remarkable experiments in which a variant of the B16 murine melanoma tumour line which is susceptible to killing by NK cells *in vitro* showed enhanced growth, faster induction time and markedly increased metastatic capability in *beige* mice, which are deficient in NK cells, compared with control mice (Talmadge *et al.* this issue of *Nature*, page 622; see also Karre *et al.* this issue of *Nature*, page 624, on the possible role of NK cells in resistance to syngeneic leukaemias in mice). Talmadge *et al.* find that induction of interferon by infection with LCM virus decreased the tumour growth rate and metastatic frequency even in *beige* mice. Treatment of *nude* mice, which ordinarily reject tumour cells persistently infected with virus, with antibodies to mouse interferon not only results in the growth of the tumours, but in their generalized metastasis (Reid, *et al. 3rd Intl Workshop on Nude Mice* in the press).

Much less is known about the role of the interferon-NK system in human disease. In diseases possibly caused by persistent viruses such as multiple sclerosis and lupus erythematosus, diminished production of leukocyte interferon and NK activity *in vitro* by peripheral lymphocytes in some patients suggest that such deficiencies could predispose to persistent virus infection (Minato *et al. op. cit.*). Several investigators are finding diminished NK activity in patients with tumours (Hersey *et al. Brit. J. Cancer* **40**, 113; 1979; Pross & Baines in *Natural Cell-Mediated Immunity against Tumours* (ed. Herberman) Academic Press, in the press). Diminished NK activity was found in familial melanoma patients and their relatives suggesting that the deficiencies in these patients may be a predisposing factor to their disease, rather than simply the result of disseminated metastases.

Questions for the future

The direct biological activity and immunoregulatory activity of interferon has raised high expectations and many questions. Yet, there are very few cases in which it has been possible to disentangle the direct effects of interferons *in vivo* on tumour cells and target cells for viruses from their effects on the immune system, and this will probably require analysis of the genetic make-up of peculiar mouse strains, for example *beige* and *nude* mice.

If a positive feedback circuit regulates the differentiation of NK cells and precursors, are there factors which normally restrict or regulate this process? One wonders particularly whether T cells, or a subset, intervene to restrict the development of NK cells. Do leukocyte and 'immune' interferons have any unique functions or target activities?

Are NK cells clonal, and do they have specificity in the immunological sense, or merely selectivity? If the latter, what are the common surface determinants which they recognize in various tumour and virus-

infected cells; if the former, can these cells be cloned and specifically amplified *in vitro*?

Is the role of interferon in natural resistance to infection or tumour spread primarily local? Can systemic administration provide an adequate effect on regional responses?

Does a deficiency either in interferon production or NK activity predispose to persistent virus infections or metastases, and can *in vitro* tests of these parameters be used as diagnostic adjuncts? In patients with these diseases does interferon augment NK activity in some patients and not others, and if so, can prescreening of patients scheduled for interferon therapy for *in vitro* sensitivity to interferon identify those showing better clinical response *in vivo*? Answers to many of these questions

are bound to influence a great deal more than the prosperity of the interferon industry.

In understanding immunological effects of interferon, it is important to note that there are very few immunology laboratories that are able to produce and purify the various types of interferon and at the same time pursue their effects on the immune system. For rapid progress, immunologists will be dependent on the good will and collaboration of the interferon experts providing them with purified material for these studies.

At a minimum, one may hope that the immunologists could repay them both by producing monoclonal antibodies against various interferons and providing insights into the mechanisms by which their product exerts its effect. □

Interstellar spectroscopy

from B.E. Turner

ALMOST a dozen years ago the first interstellar polyatomic molecules were detected by ground-based radio astronomy in the microwave region. Today, the list of known interstellar molecules stands at 53 and includes 90 different isotopically substituted species. These molecules have provided a new and unique tool for studying the dense, cool component of the interstellar medium, a component now recognized as comprising half of the interstellar medium by mass. Molecular clouds which range up to 10^7 solar masses are the sites of star formation throughout the Galaxy and have an important, if not dominant, effect on the dynamics and kinematics of both the interstellar medium and of the stellar populations immersed in it.

Both the physical conditions in these interstellar clouds, and their chemistry, have been subjects of intense study over the past decade. The molecular species identified range in complexity from 2 to 11 atoms. Not surprisingly, the more complex the molecule, the more speculative our understanding about how it is formed. Thus, meaningful tests of the current theories about chemical processes in interstellar molecular clouds necessarily involve the simpler of the known species. To be useful either as tests of chemical models or as probes of physical conditions, these simple species must have reliably determined abundances.

Adequate study of important interstellar molecules is hampered or prevented by the Earth's atmosphere as it absorbs strongly in several bands of O_2 and H_2O at wavelengths longer than $800\ \mu\text{m}$, and fairly continuously in a myriad of water transitions over the entire sub-millimetre region below $800\ \mu\text{m}$. The most important

interstellar molecules are light, and have many important transitions in the sub-millimetre region, observations of which are needed to determine abundances. Among these are CO and HD, which along with atomic carbon provide the principal cooling transitions of dense clouds. Also included are all of the basic hydrides of the hydrocarbon, nitride, and hydroxyl families and of a number of elements (Mg, P, Na, Cl, Al, Ca, Fe) not yet identified in interstellar molecules. Key transitions of interstellar species such as H_2O and O_2 are of course absorbed by the corresponding telluric lines.

In the past few years a strong effort has been mounted to study extraterrestrial lines from above the troposphere. These studies conducted with the Kuiper Airborne Observatory (operated by NASA Ames Research Center) have begun to pay handsome returns with the airborne detections of the important interstellar line of H_2O at 183 GHz, two sub-millimetre lines of CO and one of H_2O , and, very recently, the 492 GHz fine-structure transition of atomic carbon.

In a series of observations from the KAO beginning in late 1976 J. W. Waters and colleagues (*Astrophys. J.* **235**, 57; 1980) made the first detection of an interstellar molecular line inaccessible from the ground, the $3_{13}-2_{20}$ transition of H_2O at 183 GHz. Interstellar H_2O has long been observed from the ground via its $6_{16}-5_{23}$ transition at 22.2 GHz. In normally excited H_2O such as occurs in the Earth's atmosphere, the 22 GHz line is very weak, thus explaining the modest (a few percent) absorption by the atmosphere here. By contrast, the interstellar version of this line is far from normal — it is observed without exception as a powerful maser whose pumping mechanism is not yet understood so that little information about the

abundance of interstellar H_2O can be deduced. Here is where the new line discovered by Waters *et al.* comes in. In the Orion molecular cloud where the new line is seen, lines of most other molecules have a characteristic emission profile (a sharp spike superimposed upon a broad pedestal) indicative of rapid mass motions within a small energetic core surrounded by a larger more quiescent region. Molecules known to be masering in Orion (OH , 22 GHz H_2O , CH_3OH) do not show this profile, but non-masering molecules do. Although they are unable to resolve the emission region spatially with their 91 cm telescope, Waters *et al.* find that the 183 GHz water emission emulates quite well the spike-plateau profile of other molecules, and therefore conclude that this water transition is reasonably normally excited. This spike feature apparently arises in the quiescent region whose size (about 4 arc min) and temperature (about 80 K) are well determined from the lines of many other molecules. If the 183 GHz line of H_2O is thermalized at this temperature, Waters *et al.* estimate the H_2O abundance at a few times $10^{17}\ \text{cm}^{-2}$. If one calculates the excitation produced by known infrared radiation fields and by collisions with other molecules, then H_2O abundances as large as $10^{18}\ \text{cm}^{-2}$ can result. The net abundance uncertainty (a factor of 10) is in fact no greater than applies to many other molecular species observed from the ground, and will be greatly reduced by use of a larger airborne telescope which resolves the source. This first result already yields important ratios for the interstellar chemist: the fractional abundance of H_2O is about 10^{-6} , consistent with ion-molecule chemical models. From ground-based studies of HDO, one deduces that $HDO/H_2O \gg D/H$, showing that water, like several other interstellar molecules, concentrates deuterium with great efficiency.

Even more recent detections of hitherto inaccessible interstellar lines and molecules have been made from above the troposphere by T. L. Phillips and colleagues, using the KAO (*IAU Symposium 87*, Mt. Tremblant, Quebec, August 1979). The $4_{14}-3_{21}$ transition of H_2O at 380 GHz has a spectral profile in the Orion source similar to that of the 183 GHz line, and thus is consistent with the picture discussed by Waters *et al.* Carbon monoxide, the most important tracer of the dense molecular interstellar gas, is now much better understood thanks to the detections in mid-1979 of the $J=3-2$ and $4-3$ transitions at 345 and 461 GHz respectively. These higher-lying CO transitions often exhibit marked self-reversals in their profiles, which are usually not evident in the lower-lying lines previously studied from the ground, and which are indicative of complex temperature, velocity, and density variations within the clouds. Such self-reversals may also bear on the classical

B.E. Turner is at the National Radio Observatory, Charlottesville, Virginia.

question of whether molecular clouds collapse, and at what rate, a question of direct consequence to the puzzle of how stars form from such clouds. Late last year Philips and his colleagues made perhaps the most important discovery yet from the KAO, the fine-structure ($^3P_1-^3P_0$) ground-state transition of neutral atomic carbon at 492 GHz (BAAS 11, 686; 1979). The line is apparently observable everywhere that CO is seen. Not only is carbon chemistry predominant in interstellar molecular clouds, but models predict extremely abrupt transitions in the ionized, neutral, and molecular forms of carbon as a function of optical depth of the cloud. Such predictions can now be checked. Neutral atomic carbon is also expected to be one of the major coolants of clouds by emission in its two fine structure transitions. A much better determination of heating-cooling mechanisms in these clouds is now possible. It is quite probable

that the new carbon line will take its place along with the atomic hydrogen line at 21 cm and the CO line at 2.6 mm as one of the three most important interstellar lines in the radio-infrared spectrum.

In the short time that observations of the interstellar medium from above the troposphere have been possible, the results have been quite remarkable. More of the same can be expected. Yet to come are searches for the hydrides of many elements, the other fine structure transition of carbon (at 812 GHz), important molecules such as HD and O₂, and vibrationally excited molecular species of all kinds including several (such as acetylene and carbon dioxide) that because of their symmetry do not radiate at microwave frequencies. The KAO has been the vital first step in opening one of the last, and potentially richest, regions of the astronomical spectrum — that of the sub-millimetre and far infrared. □

Spinning yarns about plant cells

from Clive Lloyd

WHERE and how is cellulose made? With rare exception it doesn't seem to be made inside plant cells, neither does it appear to be spun-out from precursors away from the cell membrane. In fact, it has long been held that cellulose is synthesized at the plasma membrane but although convincing images exist of the sites of assembly in algae there is still little agreement about the nature of the synthesizing machinery in higher plants. Membrane fractions will add radioactive glucose (from nucleotide-glucose) to primer whereas unbroken protoplasts will not and so it is thought that the activated glucose must be donated to membrane-located cellulose synthetases at their cytoplasmic face. Splitting the lipid layer open by freeze-etch/freeze-fracture techniques should reveal any large trans-membranous assemblies but it is precisely this area of producing artefact-free images of putative "assemblyses" which is fraught with problems and has prolonged the debate about what happens in higher plants.

It was in 1963 that Preston suggested the existence of enzyme particle complexes in the plasma membrane responsible for the synthesis of cellulose microfibrils and they were soon found by the new technique of freeze-etch/freeze-fracture (see Preston *The Physical Biology of Plant Cell Walls*, Chapman & Hall, 1974). With the algae *Oocystis* (Brown & Montezinos *Proc. natn. Acad. Sci. U.S.A.* 73, 143; 1976) and *Glaucoecystis* (Willison & Brown *J. Cell Biol.* 78, 103; 1978) and the bacterium *Acetobacter xylinum* (Brown, Willison & Richardson *Proc. natn. Acad. Sci. U.S.A.* 73, 4565; 1976), intramembranous com-

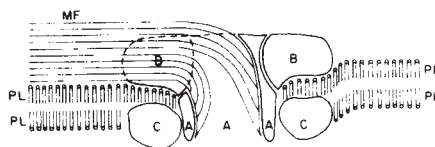


Fig. 1 A model demonstrating the proposed relationship of terminal complexes and rosettes in association with the plasma membrane. The membrane, microfibrils, and microfibril-associated structures are drawn approximately to equivalent scale. The dimensions of the rosettes (C), the terminal complex (A and B) and the microfibril (MF) are derived from freeze-fracture images. (from Mueller & Brown, *op. cit.*)

plexes revealed by this technique pack together in lines and are associated with cellulose microfibrils on the external face. Linear (as well as hexagonal) arrays of particles are also seen in fractured membranes of regenerating protoplasts from the higher plant *Skimmia japonica* (Robenek & Peveling *Planta* 136, 135; 1977) when glycerol is used. However, Davey and Mathias (*Protoplasma* 100, 85; 1979) have shown that the formation of both hexagonal and linear arrays of membrane particles can be attributed to the use of glycerol — a cryoprotectant known to induce similar artefacts in animal cells. So, do intramembranous assemblies participate in processing the higher plant cell wall or do they not? They do according to work from R.M. Brown's laboratory. In 1976, Mueller, Scott and Brown (*Science* 194, 949) presented pictures of the extracellular face of fractured membranes from untreated corn roots, showing globular complexes attached to the ends of nascent cellulose microfibrils. Willison and

Grout (*Planta* 140, 57; 1978) also avoided chemical fixation or glycerol treatment and showed similar globules attached to the ends of microfibrils in radish roots. Tomato and tobacco protoplasts were also studied but even though they were presumably regenerating a wall, no impressions of wall microfibrils were made into the membrane, thus preventing any correlation between cellulose microfibrils and the intramembranous particles. Without their walls, protoplasts are fragile and are prevented from swelling and bursting by regulating the osmotic concentration of the suspension medium with sugars. It seems likely, then, that in contrast to whole tissue, a wall in the process of being regenerated by a protoplast may be physically incapable of imprinting its image through to the membrane's fracture-plane where the particles are seen. For several reasons, therefore, it would seem that protoplasts may not be ideal for studying wall synthesis and now Wilkinson and Northcote (*J. Cell Science*, in the press) reinforce the message by reporting that even plasmolysis (a routine preliminary to preparing protoplasts from suspension cells) induces proteins to crystallize in the plane of the membrane as the area of surrounding lipid bilayer is reduced during retraction of the protoplast from the cell wall. It would seem that particle arrays produced by any treatment (not just glycerination) which causes plasmolysis could be artefactual.

This readiness of intramembranous particles to pack together under varied experimental conditions tends to sap confidence in the physiological relevance of ordered particle arrays to cellulose biosynthesis. It is satisfying therefore to find further good evidence in higher plant cells which favours the idea that specialized particle rosettes, quite unlike the more dubious hexagonal arrays, are associated with cellulose synthesis. Mueller and Brown (*J. Cell Biol.* 84, 315; 1980) have now confirmed the presence of such assemblies in untreated corn, pine and mung bean seedlings by freezing the material directly in solid nitrogen. They find that terminal complexes are associated with impressions made by cellulose microfibrils into the external leaflet of the lipid bilayer. It used to be said that cellulose formed a continuous meshwork without visible free ends but in one case a microfibril has been pulled through this leaflet but is still seen to terminate in the central subunit of such a complex. On the complementary protoplasmic side of the fracture, rosettes of membrane particles are commonly found. Both the rosettes and the terminal complexes are associated with the ends of microfibril impressions and it is envisaged that these could be complementary to one another such that the central subunit (which terminates the cellulose microfibril) normally sits within the rosette (Fig. 1). Once again, these investigators emphasize the importance of cellular

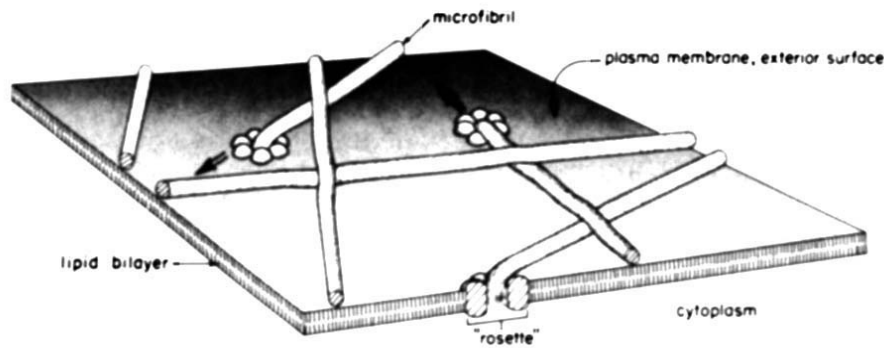


Fig 2 Model of microfibril deposition during primary wall formation in *Micrasterias*. Above: side view. Below: surface view. Single rosettes apparently give rise to randomly oriented microfibrils. (from Giddings *et al. op. cit.*)

turgor for visualizing the terminal complexes. A brief (1 min) exposure to glycerol before freezing causes the cells to lose turgor whereupon microfibril impressions are no longer observed. An accompanying paper by Giddings, Brower and Staehelin (*J. Cell Biol.* **84**, 327; 1980) proposes a similar view of cellulose biosynthesis drawn from work on the green alga *Micrasterias denticulata*. Here, the hexagonal rosette, together with the central microfibril-terminating subunit, are pictured as a transmembrane complex responsible for the elaboration of the basic single 5 nm cellulose microfibril (Fig. 2). Rows of such complexes are believed to form sets of microfibrils which aggregate laterally to form the larger fibrils composing the secondary wall. This view of wall formation in *Micrasterias* supports that of Kiermayer and Dobberstein (Protoplasma,

77, 437; 1979) who also indicated that these specialized areas of membrane are derived by the insertion of Golgi-derived 'flat vesicles' into the plasma membrane.

These investigators are properly cautious in designating these particle complexes as the cellulose synthesizing machinery. However, if correct, the next major question for plant morphogenesis is what determines the orientation of cellulose deposition and hence plant cell shape. Perhaps the linearity of crystalline cellulose fibrils is sufficient to steer through the membrane the ordered groups of particle complexes which extrude them. On the other hand, there are grounds to believe (in higher plants) that cytoplasmic microtubules, which may be bridged to the plasma membrane, influence the orientation of cellulose deposition — but that is another story. □

Interstellar violence observed

from Robert Kirshner

SUPERSONIC shocks created by stellar winds or supernova explosions are capable of heating large volumes of the interstellar gas in our Galaxy to temperatures approaching those of the solar corona. Thermal emission from this torrid interstellar medium emerges as soft X rays in the vicinity of 1 keV. The A-2 experiment on the HEAO-1 satellite spent the time from its launch in August 1977 until its demise in January 1979 in a patient effort to construct an all-sky map in soft X rays. The virtue of this approach is that it can reveal new sources in unexpected places and that it can detect very large and nearby objects that are too diffuse for imaging detectors. This scheme has paid off handsomely in the discovery of a very large (13° diameter) soft X-ray source in the constellation Cygnus which may be the fossil remains of a series of violent explosions in the interstellar gas.

As reported at a recent American Astronomical Society Meeting (San Francisco, 16 January, 1980) Webster Cash

Robert Kirshner is Assistant Professor of Astronomy at the University of Michigan.

of the University of Colorado and his collaborators from Berkeley, Caltech and the Jet Propulsion Laboratory have examined a portion of the soft X-ray survey. From their map of the Cygnus region, they find several known supernova remnants together with an immense horseshoe of emission they term an X-ray superbubble. Their data show that the emission comes from gas near 2 million K, and that there is substantial cold gas between us and the emission source which absorbs many of the X rays. From the absorption data, Cash *et al.* estimate that the distance to the giant loop of X-ray emission is about 2 kiloparsecs. Although the distance is uncertain, it implies that the diameter of the X-ray source is an astonishing 450 parsecs, 10 times larger than an ordinary supernova remnant. From the distance and the observed X-ray flux, they use the estimated temperature to derive an electron density of about 0.02 cm^{-3} , and a total thermal energy content of about $6 \times 10^{51} \text{ erg}$. This is about an order of magnitude larger than the total energy output of an ordinary supernova.

To understand this object, Cash *et al.* have considered its possible relation to other interesting objects in the same region of the galactic plane. A ring of optical filaments emitting $H\alpha$ coincides with the boundary of the X-ray ring. The HEAO-1 observers suggest that the same interstellar shock which has heated the interstellar gas to produce the X-rays may be responsible for exciting these optical filaments. A large expanding loop of neutral hydrogen, one of several reported by Carl Heiles on the basis of 21 cm observations (*Astrophys. J.* **229**, 533; 1979), lies in the same direction. The most important object in the region is the group of young stars known as the Cygnus OB2 association, which lies within the boundaries of the X-ray source.

This assembly of hot and luminous stars might provide the energy for the X-ray superbubble, either through a slow sustained energy input due to violent stellar winds or by impulsive energy input from supernova explosions of the most massive stars in the association. Cash *et al.* show that a single supernova explosion would require 10^{54} erg to create the X-ray source, an energy 10^3 times larger than any observed supernova. The energy demands need not be so severe if the supernovae are frequent enough. In that case, each successive event will take place in the low density cavity blown in the interstellar gas by its predecessor. In this way, a bubble of the immense size observed with HEAO-1 could be created by several successive supernova explosions in the OB association.

The Cygnus superbubble is not the only large region of the interstellar gas which has been observed to be filled with a hot low-density medium. Reynolds and Ogden (*Astrophys. J.* **229**, 942; 1979) have examined the energetics and kinematics of a 280 parsec diameter region which is connected to the I Orion OB association that excites the Orion Nebula. Just as in the Cygnus case, faint $H\alpha$ filaments and 21 cm features outline a large soft X-ray source. Because the Orion bubble is only 460 parsec away, this object covers a full 40° on the sky. The measurements of Reynolds and Ogden show that the optical filaments are expanding at about 20 km s^{-1} . The picture in Orion, as in Cygnus, seems to be that several stars from the OB association have already lived out their lives and detonated as supernovae, carving out a large, hot bubble in the interstellar gas from their overlapping shock waves.

The idea that a substantial fraction of the interstellar medium might be maintained at a temperature of 10^6 K and a density near 10^{-2} cm^{-3} by means of overlapping supernova remnants is not new. In 1974, Cox and Smith (*Astrophys. J. Lett.* **189**, 105; 1974) pointed out that supernova explosions are frequent enough that their remnants might often connect, and more recently McKee and Ostriker (*Astrophys. J.* **218**, 148; 1977) have developed a comprehensive model for the interstellar

medium which predicts that coronal gas, in rough pressure equilibrium with cooler phases of interstellar gas, may fill a substantial fraction of the galactic disk. The HEAO-1 survey, and particularly the new Cygnus superbubble, provide direct evidence that the mechanisms envisioned by these theorists actually do take place. Both the Cygnus object and the Orion bubble emphasize that OB associations, where large collections of massive stars can be found, may play a prominent part in creating the overlapping supernova remnants that can fill large portions of the galaxy with coronal gas. □

Topical fusion

from M. Keith Matzen

EDWARD Teller (Lawrence Livermore Laboratory) opened a recent meeting on inertial confinement fusion (ICF)* by stressing the importance of research over premature emphasis on development. We were reminded that an ICF reactor is technologically more complex than a fission reactor, that it took 20 years to develop an economically feasible fission power plant after the demonstration of the first fission reactor, and that the first demonstration of an ICF reactor is many years away. Teller believes that there is a "small chance" that a magnetic fusion reactor could be built in this century, but that the chances of an ICF reactor in the same time are "extremely remote". On the other hand he believes that a magnetic fusion-fission hybrid could be built in about 10 years and should be pursued as an energy-producing alternative.

According to Lawrence Killion (Department of Energy, Office of Inertial Fusion) the emphasis in the ICF programme during the next decade will be to define an appropriate driver. He stated that lasers are no longer considered necessarily to be the dominant long-term driver because of their high cost and the uncertainties in the laser-target coupling. The goals for the programme in the 1980s must include the demonstration of a high-gain driver-target system, the determination of the size of a reactor driver, the understanding of driver-target coupling, the achievement of breakthroughs in driver technology, and the achievement of breakthroughs in target fabrication technology.

John Nuckolls (LLL) provided further requirements and perspectives for future ICF research. The basic requirement of an ICF reactor system is an energy gain of greater than 100 with a driver energy of less than 10 MJ (10^7 Joules). Computer calculations now predict that an energy gain of 100 can be obtained with driver

energies of 2 MJ ($\pm$ a factor of 3). He stated this energy requirement had not changed in several years, in contrast to the inflation of the driver energy required for breakeven (gain 1) targets. The increase in the break-even driver energy from 1 kJ (see *Nature* 239, 139; 1972) to the present estimate of 300 kJ ($\pm$ a factor of 3) is due to the inability to approach the theoretical limits of the early calculations, both in laser technology and in laser-target coupling physics.

The importance of understanding the laser-target coupling physics can be illustrated by a number of key theoretical issues: fast electron transport, competition between Brillouin scattering and inverse bremsstrahlung absorption, production of fast electrons, and beam filamentation. He suggested that future experimental work should emphasize studies with longer pulse lengths and shorter wavelengths, measurements of the underdense plasma, studies of the fast electron production and transport, and studies of bandwidth effects. Although ions (both heavy and light) must solve problems of achieving sufficient power density on target, their target-coupling problems seem minimal. There is a rigorous upper limit for the ion range in a target and there are apparently no problems with fuel preheat from the electron and nuclear reactions. The problem of transporting ion beams to the target is difficult but seems to be soluble. In summary, Nuckolls gave the following estimates for the parameters of a driver for an ICF reactor: 3 MJ energy, 150 TW power, 5-10% efficient, beam focusable to less than 5 mm, 5 Hz repetition rate, cost less than 10^9 dollars, and operate for 10^{10} shots. The candidate reactor drivers are short wavelength lasers (KrF for example), heavy ion beams (HIB), and light ion beams (LIB). The drivers that may show target ignition (DT gain ~ 5 , target gain ~ 0.1) in the mid 1980s are Nova (at Lawrence Livermore Labs) and PBFA (at Sandia National Labs).

The recent emphasis in experimental programmes has been in the areas of intermediate density targets (10 times liquid DT density), techniques for measuring these compressed densities, X-ray backlighting techniques, ablative acceleration of thin foils, characterizing the back-reflected laser light, characterizing the thermal conductivity and fast-electron production by using multiple-layer targets, measuring the driver-target coupling efficiency, and experiments to examine the stability of implosion systems. With the trend towards shorter wavelength laser-drivers, several laboratories have begun experiments with doubled, tripled, and quadrupled glass laser light (wavelengths of 0.53, 0.35, and 0.26 μm , respectively). Although most experiments are still in their initial stages, the early results generally show better coupling to the target, decreased fast electron production, and better thermal conduction

as the laser wavelength decreases. However, Fred Mayer (KMS Fusion) presented measurements in which the same absorption was observed for 0.53 μm and 1.06 μm irradiation of spherical targets with 60 to 100 ps pulses. These results are in contrast to the strong wavelength dependence observed in 0.26 μm , 0.53 μm , and 1.06 μm irradiation of planar targets (Fabre *et al.*, Ecole Polytechnique Palaiseau, France). The near future will see a great deal of effort devoted to short-wavelength ($\leq 0.53 \mu\text{m}$) irradiation experiments.

ICF reactor designs are still in the conceptual stages. Protection of the reactor walls from the radiation environment is a major consideration. Several reactor designs have been considered, although flowing-liquid-metal-walled reactors have received most attention and represent the only design in which protection is provided against all the fusion reaction products. Jerry Kulcinski (University of Wisconsin) reported that work is needed in the areas of fusion reactor economics, safety and environmental impact, as well as in recovery of pellet debris, protection of final focusing elements, small LIB reactors, minimum sized HIB reactors, and pulsed neutron damage. Both Kulcinski and Noel Amherd (EPRI) stated that LIB technology appears to offer the only approach to small ICF reactors at this time. John Caird (Bechtel National, Inc.) has projected the capital cost of a HIB power plant to be from 4.2 to 6.6 billion dollars for plants in the range from 1.125 to 2.25×10^9 Watts-electrical.

With the trend towards larger, more complicated high-gain targets, the technology of target fabrication continues to be challenged. High-gain targets, in general, consist of multiple shells and larger aspect ratios (ratio of spherical shell radius to shell thickness). Thus the questions of target-shell finish and stability of implosion become more important. The technology for building these multishell targets is not yet available. At this meeting, the emphasis was on target-shell and fuel-fill characterization techniques, pellet fabrication methods and mass production of pellets. Novel approaches to target fabrication include using a low-gravity environment to form large-diameter spherical bubbles, laser drilling of 2 to 4 μm holes in microballoons, followed by gas filling and hole plugging, and structural modification of polymers to reduce crystallinity and improve the surface finish. Nondestructive fuel fill characterizations is another area of active research.

Progress on the construction of new, larger drivers continues. Nova, the 100 kJ, 1-3 ns glass facility at Lawrence Livermore Laboratory should be available for target

M Keith Matzen is in the Laser Theory Division of Sandia Laboratories, Albuquerque, New Mexico.

*The topical meeting on Inertial Confinement Fusion (ICF) was held in conjunction with the Conference on Laser and Electro-Optical Systems (CLEOS) in San Diego, California, February 26-28, 1980.

experiments in 1984. An interesting sidenote is that Shiva was operational 1 month after the earthquake in Livermore and suffered only minor damage to the optical system. At Los Alamos, the current scenario for the Antares CO₂ laser facility envisages the construction of two of the previously proposed six beam lines so that 40 kJ of CO₂ laser energy can be provided for target experiments in 1983. Recent improvements in the triple-pass amplifiers of the Helios system at Los Alamos will increase their energy on target from 5 to 10 kJ by early summer. The light-ion-beam facility at Sandia Laboratories (PBFA-1) is scheduled for completion this year and will provide 1 MJ and 30 TW. The upgrade of this facility, PBFA-II, will deliver 3.5 MJ at 100 TW and will be available for experiments in 1986. A 10-kJ heavy-ion-beam target physics facility has been proposed for construction and operation by 1984. As mentioned previously the programme for new laser development is emphasizing short wavelength lasers. The present candidates are the rare gas halide lasers KrF, XeCl, and XeF. All these require a pulse compression technique to obtain pulses in the neighbourhood of 20 ns. Papers at this meeting showed the feasibility of both Raman compression and pulse stacking (angular multiplexing) to achieve shorter pulses. A major drawback of these short-wavelength systems is the low damage threshold of the optical components and the resulting high cost of the optical system. This problem has not been adequately addressed. Since the tradeoffs between these systems will depend very strongly on the results of future laser-target coupling experiments, the option to use longer-wavelength lasers has not been eliminated. □

Thunderstorms and substorms: any connection?

from D. J. Southwood

It is a familiar fact that substantial electric fields are produced in the atmosphere during a thunderstorm. As well as a steady (d.c.) field being present, thunderstorm lightning flashes are a source of transient a.c. electromagnetic signals over a wide frequency range. A less familiar electromagnetic storm phenomenon is the magnetospheric substorm, a disturbance of the ionosphere and magnetosphere in the tenuous outer ionized layers of the Earth's atmosphere. The substorm's most dramatic manifestation is in intense

auroral displays. The aurora results from the strong electric coupling between the magnetosphere and the ionosphere below, a process which can require strong currents to flow between them. The aurora occurs when the currents are carried by energetic electron beams which give rise to light as they strike the denser lower ionosphere.

The ionosphere and magnetosphere are very good electrical conductors while the troposphere, the lowest layer of the atmosphere and the one in which thunderstorms occur, is insulating. There have been few attempts to consider whether there could be any significant electrical coupling between the troposphere and the outermost regions except by those workers (Markson *Nature* 273, 103; 1978; Herman & Goldberg *J. Atmos. Terr. Phys.* 40, 121; 1978) looking at "sun-weather relationships". The search is for causal links between activity in the tenuous outer layers of the atmosphere and the much denser bottom 5-10 km where the weather occurs. Any such connection must be subtle. The energy associated with tropospheric motions far exceeds that associated with any magnetospheric effect. Electric coupling (Markson *op. cit.*) is one possible connection.

Data from a recently reported balloon flight (Bering *et al. J. geophys. Res.* 85, 55; 1980) in northern Canada that took place during the simultaneous occurrence of a thunderstorm and a magnetospheric substorm has permitted a direct look at some interrelations between these apparently disparate disturbances in very different regions. The balloon was some 10-20 km above the top of the thunderstorm activity (at ~ 10 km) but some 60 to 70 km below the bottom of the E region ionosphere. The thunderstorm was localised and centred some 30-40 km north-west of the balloon launch site in Roberval, Quebec. In contrast the substorm was a worldwide event. Magnetograms from the USSR, Alaska, Canada and Greenland showed that magnetic perturbations due to the magnetospheric substorm were present all around the auroral zone. The balloon was instrumented to measure vertical and horizontal electric fields, electrical conductivity and X-ray emissions. The latter provide information on energetic electrons precipitated from the magnetosphere but which are stopped at much higher altitudes than the balloon's flight height. On the ground VLF (very low frequency) radio receivers, goniometers and other equipment for support measurements were deployed.

The d.c. electric field at the balloon was dominated by the presence of the thunderstorm. There are uncertainties in extrapolating the observed field upwards to the ionosphere. On this flight a reasonable estimate suggests the thunderstorm charge distribution should produce a field of 10^{-7} V m⁻¹ in the ionosphere. In contrast the detection of

echoes by the Cornell University auroral radar which was operating during the flight shows the actual ionospheric field must have exceeded 10^{-2} V m⁻¹. It seems safe to say the thunderstorm electric field had no effect on the magnetosphere and ionosphere.

Thunderstorms are a strong source of a.c. signals in the atmosphere and these can penetrate upwards far more effectively. Lightning flashes associated with the storm generate ULF (ultra low frequency) and VLF electromagnetic noise. In the 1-30 kHz VLF band lightning-flashes have long been known as a source of whistler mode plasma waves in the magnetosphere. A complicated interaction between magnetosphere and atmosphere can then be envisaged because lightning-launched whistler mode signals can interact with the magnetospheric energetic electron population. In particular, waves can scatter electrons onto orbits which bring them down into the atmosphere or lower ionosphere along the Earth's magnetic field. Here they experience collisions and in particular generate secondary X rays. The X-ray emissions seen on the balloon flight were of the microburst type, short bursts (< 1 s) which characteristically occur during substorms. A key question thus was, were any microbursts caused by lightning-induced VLF? The authors argue there were. Their test is to look at how many microbursts occur within 1.5 s of a lightning flash. 62 out of 95 nearby flashes appeared to have closely correlated bursts of electron precipitation following them. The authors estimate a 1% probability of this being chance.

The flashes also produce substantial power in the much lower ULF band (< 1 Hz). They conclude this could have a substantial effect in the ionosphere and suggest the thunderstorm-induced ionospheric ULF field could considerably exceed the substorm-induced ULF fields in the ionosphere. Whether this is of significance in magnetospheric terms is another matter. No measurements were made of magnetospheric parameters that could have been directly affected (such as ion precipitation).

The authors' most surprising measurement is that of the local electrical conductivity. It is about half its fair weather value. This certainly contradicts electrical coupling mechanisms between atmosphere and magnetosphere which invoke a substorm-induced conductivity enhancement as a trigger in a coupling mechanism.

All in all the report cuts down our ability to speculate on electric coupling of magnetosphere and atmosphere. D.C. coupling seems unimportant. The conductivity measurement is a puzzle. It does seem ULF and VLF phenomena can couple but such coupling should not be of dynamic significance in either region. The hard fact is we seem no nearer a substantive connection. □

Distributed feedback lasers

from R.G. Harrison

SINCE the gain and so the output power of most gas lasers increases as the inner diameter of the laser tube decreases, considerable attention has been given to the design of laser systems with very narrow bores. For these lasers the electromagnetic resonator modes are altered in a fundamental way from those obtained from conventional open resonator cavities formed by mirrors; the radiation field distribution being no longer determined by the mirrors but by the modes of the hollow waveguide. (Marcatili & Schmetzer, *Bell Syst. Tech. J.*, **43**, 1783; 1964). Laser oscillation is normally sustained in such systems by feedback to the tube from external mirrors. However, an alternative method recently demonstrated is that of distributed feedback (DFB) by which a forward propagating mode is coupled to its corresponding backward mode through scattering from periodic diameter or width changes in the tube. The distinctive feature of this coupling is that the scattered wave excited at a certain position in the waveguide is added in phase to the already existing wave, efficient coupling occurring only when the frequency of the ripple structure in the tube is nearly equal to twice the propagation constant of the lasing mode. Since it is not necessary that this structure be sinusoidal a convenient alternative can be provided by rectangular grooves, coupling then being provided by the fundamental frequency of the periodic rectangular ripple.

For distributed feedback systems the threshold for oscillation increases as the order of resonance of the resonant mode increases. Thus they offer the advantage of operation on the lowest order waveguide mode at a single frequency determined by the resonant mode q ; discrimination from other modes being controlled by choosing the gain of the laser such that all higher order modes than the lowest one suffer overall loss. This together with low laser thresholds and the requirement of no external optics makes the technique potentially attractive. Furthermore, the method should not be constrained by the material or cross-sectional geometry of the light tube, and as such it may be applied to many gas laser systems. These include optically pumped lasers for the near to far infrared using either dielectric or metal tubing and also transverse electrically excited high pressure systems and Stark tuned optically pumped lasers, both comprising parallel flat electrodes separated by dielectric spacers to form a rectangular section, metal-dielectric, waveguide.

Distributed feedback laser action was first reported some years back by Kogelnik

and Shank (*Appl. Phys. Letts.*, **8**, 152; 1971) for a dye system. However it was only recently that such action was achieved in a gas system (Affolter & Kneubuhl, *Phys. Lett.*, **74A**, 407; 1979) even though the theoretical foundations for such operations had been well established (Marcuse, *IEEE J. Quantum Electr.*, **QE-10**, 413, 1974; Kneubuhl, *J. Opt. Soc. Am.*, **11**, 1; 1976). Part of the reason lies in matching the laser emission with the periodic structure of the DFB waveguide. For both dye and solid state lasers there is no problem since the emission linewidth are relatively large. However, for both optically pumped and electrically excited gas lasers, usually operated at low pressures, this is not so and matching is correspondingly critical. In the preliminary work of Affolter and Kneubuhl this difficulty has been overcome by temperature tuning the periodic structure of a CH_3F optically pumped waveguide system over a temperature range between 30 °C and 90 °C, the temperature variation of 10 °C corresponding to a relative change of $\sim 10^{-4}$ of the waveguide period. The bottom of the waveguide was periodically corrugated with rectangular grooves of depth 124 μm and period 248 μm ,

R.G. Harrison is in the Department of Physics, Heriot-Watt University, Edinburgh.

corresponding to a quarter and a half of the wavelength of the CH_3F emission and feedback was optimized by varying the height of the waveguide.

Evidence of DFB action was demonstrated in the resonant or mode features of the output energy of the CH_3F 496 μm emission as a function of the waveguide temperature. Single mode emission was obtained with optimum output at the centre of the temperature range. For higher pump energies oscillation of two modes was observed, peaked at temperatures separated by around 45 °C, corresponding to a mode separation of 450 MHz and in agreement with DFB theory. The relatively broad temperature bandwidth of around 20 °C (FWHH) over which oscillation occurs on a single mode, together with the readily accessible temperature range used in the tuning, are both particularly attractive features of the method with regard to control and operation of the system.

Whether the distributed feedback technique will eventually replace conventional cavity optics in some laser systems remains to be seen. Nevertheless these new results are promising and as such may provide the key to an accelerated effort into the research and development of such systems in the near future. □



100 years ago



From *Nature* 586, 15 April, 1880.

In *Nature*, vol. xxi. p. III, we described the ingenious planetarium recently invented by Signor Perini, and which has cost him seven years' constant labour. To-day we are able to present an illustration of this invention, which may give those of our readers who have not seen the original, some idea of its construction. The visitors are supposed to be standing underneath the dome, from which are suspended the sun and planets.

REVIEW ARTICLES

Selfish genes, the phenotype paradigm and genome evolution

W. Ford Doolittle & Carmen Sapienza

Department of Biochemistry, Dalhousie University, Halifax, Nova Scotia, Canada B3H 4H7

Natural selection operating within genomes will inevitably result in the appearance of DNAs with no phenotypic expression whose only 'function' is survival within genomes. Prokaryotic transposable elements and eukaryotic middle-repetitive sequences can be seen as such DNAs, and thus no phenotypic or evolutionary function need be assigned to them.

THE assertion that organisms are simply DNA's way of producing more DNA has been made so often that it is hard to remember who made it first. Certainly, Dawkins has provided the most forceful and uncompromising recent statement of this position, as well as of the position that it is the gene, and not the individual or the population, upon which natural selection acts¹. Although we may thus view genes and DNA as essentially 'selfish', most of us are, nevertheless, wedded to what we will call here the 'phenotype paradigm'—the notion that the major and perhaps only way in which a gene can ensure its own perpetuation is by ensuring the perpetuation of the organism it inhabits. Even genes such as the segregation-distorter locus of *Drosophila*², 'hitch-hiking' mutator genes in *Escherichia coli*^{3,4} and genes for parthenogenetic reproduction in many species⁴—which are so 'selfish' as to promote their own spread through a population at the ultimate expense of the evolutionary fitness of that population—are seen to operate through phenotype.

The phenotype paradigm underlies attempts to explain genome structure. There is a hierarchy of types of explanations we use in efforts to rationalize, in neo-darwinian terms, DNA sequences which do not code for protein. Untranslated messenger RNA sequences which precede, follow or interrupt protein-coding sequences are often assigned a phenotypic role in regulating messenger RNA maturation, transport or translation⁵⁻⁷. Portions of transcripts discarded in processing are considered to be required for processing⁸. Non-transcribed DNA, and in particular repetitive sequences, are thought of as regulatory or somehow essential to chromosome structure or pairing⁹⁻¹¹. When all attempts to assign to a given sequence or class of DNA functions of immediate phenotypic benefit to the organism fail, we resort to evolutionary explanations. The DNA is there because it facilitates genetic rearrangements which increase evolutionary versatility (and hence long-term phenotypic benefit)¹²⁻¹⁷, or because it is a repository from which new functional sequences can be recruited^{18,19} or, at worst, because it is the yet-to-be eliminated by-product of past chromosomal rearrangements of evolutionary significance^{9,19}.

Such interpretations of DNA structure are very often demonstrably correct; molecular biology would not otherwise be so fruitful. However, the phenotype paradigm is almost tautological; natural selection operates on DNA through organismal phenotype, so DNA structure must be of immediate or long-term (evolutionary) phenotypic benefit, even when we

cannot show how. As Gould and Lewontin note, 'the rejection of one adaptive story usually leads to its replacement by another, rather than to a suspicion that a different kind of explanation might be required. Since the range of adaptive stories is as wide as our minds are fertile, new stories can always be postulated' (ref. 20).

Non-phenotypic selection

What we propose here is that there are classes of DNA for which a 'different kind of explanation' may well be required. Natural selection does not operate on DNA only through organismal phenotype. Cells themselves are environments in which DNA sequences can replicate, mutate and so evolve²¹. Although DNA sequences which contribute to organismal phenotypic fitness or evolutionary adaptability indirectly increase their own chances of preservation, and may be maintained by classical phenotypic selection, the only selection pressure which DNAs experience directly is the pressure to survive within cells. If there are ways in which mutation can increase the probability of survival within cells without effect on organismal phenotype, then sequences whose only 'function' is self-preservation will inevitably arise and be maintained by what we call 'non-phenotypic selection'. Furthermore, if it can be shown that a given gene (region of DNA) or class of genes (regions) has evolved a strategy which increases its probability of survival within cells, then no additional (phenotypic) explanation for its origin or continued existence is required.

This proposal is not altogether new; Dawkins¹, Crick⁶ and Bodmer²² have briefly alluded to it.

However, there has been no systematic attempt to describe elements of prokaryotic and eukaryotic genomes as products of non-phenotypic selection whose primary and often only function is self-preservation.

Transposable elements in prokaryotes as selfish DNA

Insertion sequences and transposons can in general be inserted into a large number of chromosomal (or plasmid) sites, can be excised precisely or imprecisely and can engender deletions

or inversions in neighbouring chromosomal (or plasmid) DNAs¹²⁻¹⁶. These behaviours and, at least in some cases, the genetic information for the enzymatic machinery involved, must be inherent in the primary sequences of the transposable elements themselves, which are usually tightly conserved^{12-16,23}. Most speculations on the function of transposable elements concentrate on the role these may have, through chromosomal rearrangements and the modular assembly of different functional units, in promoting the evolution of plasmid and bacterial chromosomes, and thus in promoting long-term phenotypic fitness¹²⁻¹⁶. Most assume that it is for just such functions that natural selection has fashioned these unusual nucleic acid sequences.

Although transposable elements may well be beneficially involved in prokaryotic evolution, there are two reasons to doubt that they arose or are maintained by selection pressures for such evolutionary functions.

First, DNAs without immediate phenotypic benefit are of no immediate selective advantage to their possessor. Excess DNA should represent an energetic burden^{24,25}, and some of the activities of transposable elements are frankly destructive¹²⁻¹⁶. Evolution is not anticipatory; structures do not evolve because they might later prove useful. The selective advantage represented by evolutionary adaptability seems far too remote to ensure the maintenance, let alone to direct the formation, of the DNA sequences and/or enzymatic machinery involved. A formally identical theoretical difficulty plagues our understanding of the origin of sexual reproduction, even though this process may now clearly be evolutionarily advantageous^{1,4}.

Second, transposability itself ensures the survival of the transposed element, regardless of effects on organismal phenotype or evolutionary adaptability (unless these are sufficiently negative). Thus, no other explanation for the origin and maintenance of transposable elements is necessary. A single copy of a DNA sequence of no phenotypic benefit to the host risks deletion, but a sequence which spawns copies of itself elsewhere in the genome can only be eradicated by simultaneous multiple deletions. Simple translocation (removal from one site and insertion into another) does not provide such insurance against deletion. It is significant that recent models for transposition require retention of the parental sequence copy^{26,27}, and that bacterial insertion sequences are characteristically present in several copies per genome¹⁶. The assumption that transposable elements are maintained by selection acting on the cell does not require that they show these characteristics. The evolutionary behaviour of individual copies of transposable elements within the environment represented by a bacterial genome and its descendants can be understood in the same terms as organismal evolution. Replicate copies of a given element may diverge in sequence, but at least those features of sequence required for transposition will be maintained by (non-phenotypic) selection; copies which can no longer be translocated will eventually suffer elimination. Some divergent copies may be more readily transposed; these will increase in frequency at the expense of others. Transposable elements which depend on host functions run the risk that host mutants will no longer transpose them; it is significant that at least some transposition-specific functions are known to be coded for by the transposable elements themselves²⁶⁻²⁹. It is not to the advantage of a transposable element coding for such functions to promote the transposition of unrelated elements; the fact that given transposable elements generate flanking repeats^{16,30} of chromosomal DNAs of sizes characteristic to them (that is, 5, 9 or 11-12 base pairs) may indicate such a specificity in transposition mechanism. It is to the advantage of any transposable element to acquire genes which allow independent replication (to become a plasmid), promote host mating (to become a self-transmissible plasmid) or promote non-conjugational transmission (to become a phage like Mu).

It is certainly not novel to suggest that prokaryotic transposable elements behave in these ways, or to suggest that more frankly autonomous entities like phages have arisen from

them^{12-16,31}. However, we think it has not been sufficiently emphasized that non-phenotypic selection may inevitably give rise to transposable elements and that no phenotypic rationale for their origin and continued existence is thus required.

Transposable elements in eukaryotes

There has long been genetic evidence for the existence in eukaryotic genomes of transposable elements affecting phenotype³². These have been assigned roles in the regulation of eukaryotic gene expression and in evolution, but would have escaped genetic detection had they not had phenotypic effect. More recent evidence for transposable elements whose effects are not readily identified genetically has come fortuitously from studies of cloned eukaryotic DNAs. For instance, the Ty-1 element of yeast (which has no known phenotypic function) is flanked by direct repeats (like some prokaryotic transposons) and is transposable³³. It is present in some 35 dispersed copies and comprises some 2% of the yeast genome (like a higher-eukaryotic middle-repetitive DNA). The directly repeated δ -sequence elements flanking it are found at still other sites (just as prokaryotic insertion sequences can be found flanking transposons or independently elsewhere in the genome). Cameron *et al.* suggest that 'Ty-1 may be a nonviral "parasitic" DNA' but then go on to suggest, we think unnecessarily, that transposition 'allows adaptation of a particular cell to a new environment' (ref. 33). The repetitive elements 412, copia and 297 of *Drosophila* are physically similar to Ty-1 (and to bacterial transposable elements) and are transposable³⁴⁻³⁷. Strobel *et al.* suggest 'it is possible that the sole function of these elements is to promote genetic variability, and that their gene products may only be necessary for the maintenance and mobility of the elements themselves, rather than for other cellular processes' (ref. 37). But if maintenance and mobility mechanisms exist, then no cellular function at all need be postulated.

A large fraction of many eukaryotic genomes consists of middle-repetitive DNAs, and the variety and patterns of their interspersions with unique sequence DNA make no particular phylogenetic or phenotypically functional sense. Britten, Davidson and collaborators have elaborated models which ascribe regulatory functions to middle-repetitive DNAs, and evolutionary advantage (in terms of adaptability) to the quantitative and qualitative changes in middle-repetitive DNA content observed even between closely related species^{17,38-40}. Middle-repetitive DNAs are more conserved in sequence during evolution than are unique-sequence DNAs not coding for protein, and Klein *et al.* suggest that 'restraint on repetitive sequence divergence, either within the repeat families of a given species, or over evolutionary time spanning the emergence of different species, could be due to [phenotypic] selective pressures which prevent free sequence change in a large fraction of the repeat family members. Or perhaps repetitive sequences diverge as rapidly as do other sequences, but the type sequence of the family is preserved by frequent remultiplication of the "correct" surviving sequences' (ref. 41). The evidence for a phenotypically functional role for middle-repetitive sequences remains dishearteningly weak⁴⁰⁻⁴³, and if the calculations of Kimura⁴⁴ and Salser and Isaacson⁴⁵ are correct, middle-repetitive DNAs together comprise too large a fraction of most eukaryotic genomes to be kept accurate by darwinian selection operating on organismal phenotype. The most plausible form of "remultiplication of the 'correct' surviving sequences" is transposition. If we assume middle-repetitive DNAs in general to be transposable elements or degenerate (and no longer transposable and ultimately to be eliminated) descendants of such elements, then the observed spectra of sequence divergence within families and changes in middle-repetitive DNA family sequence and abundance can all be explained as the result of non-phenotypic selection within genomes. No cellular function at all is required to explain either the behaviour or the persistence of middle-repetitive sequences as a class.

The rest of the eukaryotic genome

Middle-repetitive DNA can comprise more than 30% of the genome of a eukaryotic cell⁴⁶. Another 1–40% consists of simple reiterated sequences whose functions remain unclear¹⁰, and Smith has argued that 'a pattern of tandem repeats is the natural state of DNA whose sequence is not maintained by selection' (ref. 47). Even unique-sequence eukaryotic DNA consists in large part of elements which do not seem to be constrained by phenotypic selection pressures⁴⁵. Some authors have argued that the intervening sequences which interrupt many eukaryotic structural genes are insertion sequence-like elements^{6,48,49}. If they are, they are likely to be the degenerate and no-longer-transposable descendants of transposable sequences whose insertion was rendered non-lethal by pre-existing cellular RNA:RNA splicing mechanisms. Such elements, once inserted, are relatively immune to deletion (since only very precise deletion can be non-lethal), and need retain only those sequence components required for RNA splicing. The rest of the element is free to drift and one expects (and observes) that only the position and number of intervening sequences in a family of homologous genes remain constant during evolution. Although evolutionary and regulatory phenotypic functions have been ascribed to intervening sequences^{6,49–51}, it is unnecessary to postulate any cellular function at all if these elements are indeed degenerate transposable elements arising initially from non-phenotypic selection. Another explanation for the origin and continued existence of intervening sequences, which also does not require phenotypically or evolutionarily advantageous roles, has been suggested elsewhere^{50,51}.

Why do prokaryotes and eukaryotes differ?

It is generally believed that prokaryotic genomes consist almost entirely of unique-sequence DNA maintained by phenotypic selection, whereas the possession of 'excess' unique and repetitive DNA sequences whose presence is at least difficult to rationalize in phenotypic terms is characteristic of eukaryotes. However, it is more accurate to say that there is a continuum of excess DNA contents; at least 1% of the *E. coli* genome can be made up of copies of six identified insertion sequences alone¹⁶. Yeast, whose genome is no larger than that of some prokaryotes, has few repeated sequences other than those coding for stable RNAs, and *Aspergillus* may have none^{52,53}. There is in general (but with many exceptions) a positive correlation between excess DNA content, genome size and what we anthropocentrically perceive as 'evolutionary advancement'. Many interpret

this as the cause and/or consequence of the increasing phenotypic complexity which characterizes organismal evolution, and attribute to excess DNA a positive role in the evolutionary process^{17–19,40}. The interplay of phenotypic and non-phenotypic forces, and the importance of understanding both in attempts to restore the 'C-value paradox' are discussed more thoroughly by Orgel and Crick in the following article.⁵⁴

There is another, simpler and perhaps obvious explanation. Non-phenotypic selection produces excess DNA, and excess DNA logically must be an energetic burden; phenotypic selection should favour its elimination^{24,25}. The amount of excess (and hence total) DNA in an organism should be loosely determined by the relative intensities of the two opposing sorts of selection. The intensity of non-phenotypic pressure on DNA to survive even without function should be independent of organismal physiology. The intensity of phenotypic selection pressure to eliminate excess DNA is not, this being greatest in organisms for which DNA replication comprises the greatest fraction of total energy expenditure. Prokaryotes in general are smaller and replicate themselves and their DNA more often than eukaryotes (especially complex multicellular eukaryotes). Phenotypic selection pressure for small 'streamlined' prokaryotic genomes with little excess DNA may be very strong.

Necessary and unnecessary explanations

We do not deny that prokaryotic transposable elements or repetitive and unique-sequence DNAs not coding for protein in eukaryotes may have roles of immediate phenotypic benefit to the organism. Nor do we deny roles for these elements in the evolutionary process. We do question the almost automatic invocation of such roles for DNAs whose function is not obvious, when another and perhaps simpler explanation for their origin and maintenance is possible. It is inevitable that natural selection of the special sort we call non-phenotypic will favour the development within genomes of DNAs whose only 'function' is survival within genomes. When a given DNA, or class of DNAs, of unproven phenotypic function can be shown to have evolved a strategy (such as transposition) which ensures its genomic survival, then no other explanation for its existence is necessary. The search for other explanations may prove, if not intellectually sterile, ultimately futile.

We thank L. Bonen, R. M. MacKay and M. Schnare for help in development of the ideas presented here, and C. W. Helleiner, M. W. Gray, C. Stuttard, R. Singer, S. D. Wainwright and E. Butz for critical discussions.

We are especially grateful to C.E. Orgel and F.H.C. Crick for discussing with us the ideas presented in the following article before publication and for encouragement and support.

1. Dawkins, R. *The Selfish Gene* (Oxford University Press, 1976).
2. Crow, J. F. *Scient. Am.* **240**, 134–146 (1979).
3. Cox, E. C. & Gibson, T. C. *Genetics* **77**, 169–184 (1974).
4. Maynard Smith, J. *The Evolution of Sex* (Cambridge University Press, 1978).
5. Darnell, J. E. *Prog. Nucleic Acid Res. molec. Biol.* **22**, 327–353 (1978).
6. Crick, F. H. C. *Science* **204**, 264–271 (1979).
7. Murray, V. & Holliday, R. *FEBS Lett.* **106**, 5–7 (1979).
8. Sogin, M. L., Pace, B. & Pace, N. R. *J. biol. Chem.* **252**, 1350–1357 (1977).
9. Fedoroff, N. *Cell* **16**, 697–710 (1979).
10. John, B. & Miklos, G. L. G. *Int. Rev. Cytol.* **58**, 1–114 (1979).
11. Zuckerkandl, E. *J. molec. Evol.* **9**, 73–122 (1976).
12. Cohen, S. N. *Nature* **263**, 731–738 (1976).
13. Starlinger, P. & Saedler, H. *Curr. Topics Microbiol. Immun.* **75**, 111–152 (1976).
14. Nevers, P. & Saedler, H. *Nature* **268**, 109–115 (1977).
15. Kleckner, N. *Cell* **11**, 11–23 (1977).
16. Kopecko, D. in *Plasmids and Transposons: Environmental Effects and Maintenance Mechanisms* (eds Stuttard, C. & Roze, K. R.) (Academic, New York, in the press).
17. Britten, R. J. & Davidson, E. H. *Q. Rev. Biol.* **46**, 111–138 (1971).
18. Ohno, S. *Evolution by Gene Duplication* (Springer, New York, 1970).
19. Hinegardner, R. in *Molecular Evolution* (ed. Ayala, F. J.) 160–199 (Sinauer, Sunderland, 1976).
20. Gould, S. J. & Lewontin, R. C. *Proc. R. Soc. B* **205**, 581–598 (1979).
21. Orgel, L. *Proc. R. Soc. B* **205**, 435–442 (1979).
22. Bodmer, W. in *Human Genetics: Possibilities and Realities*, 41–42 (Excerpta Medica, Amsterdam, 1979).
23. Johnsson, L. *Molec. gen. Genet.* **169**, 213–218 (1979).
24. Zamenhof, S. & Eichorn, H. H. *Nature* **216**, 456–458 (1967).
25. Koch, A. L. *Genetics* **72**, 297–316 (1972).
26. Shapiro, J. A. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1933–1937 (1979).
27. Arthur, A. & Sherratt, D. *Molec. gen. Genet.* **175**, 267–274 (1979).
28. Gill, R., Heffron, F., Dougan, G. & Falkow, S. *J. Bact.* **136**, 742–756 (1978).
29. MacHattie, L. A. & Shapiro, J. A. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1490–1494 (1979).
30. Huberman, P., Klaier, R., Kühn, S. & Starlinger, P. *Molec. gen. Genet.* **175**, 369–373 (1979).
31. Campbell, A. in *Biological Regulation and Development* Vol. 1 (ed. Goldberger, R. F.) 19–55 (Plenum, New York, 1979).
32. McClintock, B. *Cold Spring Harb. Symp. quant. Biol.* **16**, 13–47 (1952).
33. Cameron, J. R., Loh, E. Y. & Davis, R. W. *Cell* **16**, 739–751 (1979).
34. Finnegan, D. J., Rubin, G. M., Young, M. W. & Hogness, D. S. *Cold Spring Harb. Symp. quant. Biol.* **42**, 1053–1063 (1977).
35. Carlson, M. & Brutlag, D. *Cell* **15**, 733–742 (1978).
36. Potter, S. S., Borein, W. J. Jr, Dunsmuir, P. & Rubin, G. M. *Cell* **17**, 415–427 (1979).
37. Strobel, E., Dunsmuir, P. & Rubin, G. M. *Cell* **17**, 429–439 (1979).
38. Britten, R. J. & Davidson, E. H. *Science* **165**, 349–357 (1969).
39. Davidson, E. H., Klein, W. H. & Britten, R. J. *Dev. Biol.* **55**, 69–84 (1977).
40. Davidson, E. H. & Britten, R. J. *Science* **204**, 1052–1059 (1979).
41. Klein, W. H. *et al. Cell* **14**, 889–900 (1978).
42. Sheller, R. H., Constantini, F. D., Dozowski, M. R., Britten, R. J. & Davidson, E. H. *Cell* **15**, 189–203 (1978).
43. Kuroiwa, A. & Natori, S. *Nucleic Acids Res.* **7**, 751–754 (1979).
44. Kimura, M. *Nature* **217**, 624–626 (1968).
45. Salser, W. & Isaacson, J. S. *Prog. Nucleic Acids Res. molec. Biol.* **19**, 205–220 (1976).
46. Lewin, B. *Cell* **4**, 77–93 (1975).
47. Smith, G. P. *Science* **191**, 528–534 (1976).
48. Tsujimoto, Y. G. & Suzuki, Y. *Cell* **18**, 591–600 (1979).
49. Gilbert, W. *Nature* **271**, 501 (1978).
50. Doolittle, W. F. *Nature* **272**, 581–582 (1978).
51. Darnell, J. E. *Science* **202**, 1257–1260 (1978).
52. Roberts, T. M., Lauer, G. D. & Klotz, L. *CRC Crit. Rev. Biochem.* **3**, 349–451 (1976).
53. Timberlake, W. F. *Science* **202**, 973–974 (1978).
54. Orgel, G. E. & Crick, F. H. C. *Nature* **284**, 604–607 (1980).

Selfish DNA: the ultimate parasite

L. E. Orgel & F. H. C. Crick

The Salk Institute, 10010 N. Torrey Pines Road, La Jolla, California 92037

The DNA of higher organisms usually falls into two classes, one specific and the other comparatively nonspecific. It seems plausible that most of the latter originated by the spreading of sequences which had little or no effect on the phenotype. We examine this idea from the point of view of the natural selection of preferred replicators within the genome.

THE object of this short review is to make widely known the idea of selfish DNA. A piece of selfish DNA, in its purest form, has two distinct properties:

- (1) It arises when a DNA sequence spreads by forming additional copies of itself within the genome.
- (2) It makes no specific contribution to the phenotype.

This idea is not new. We have not attempted to trace it back to its roots. It is sketched briefly but clearly by Dawkins¹ in his book *The Selfish Gene* (page 47). The extended discussion (pages 39–45) after P. M. B. Walker's article² in the CIBA volume based on a Symposium on Human Genetics held in June 1978 shows that it was at that time already familiar to Bodmer, Fincham and one of us. That discussion referred specifically to repetitive DNA because that was the topic of Walker's article, but we shall use the term selfish DNA in a wider sense, so that it can refer not only to obviously repetitive DNA but also to certain other DNA sequences which appear to have little or no function, such as much of the DNA in the introns of genes and parts of the DNA sequences between genes. The catch-phrase 'selfish DNA' has already been mentioned briefly on two occasions^{3,4}. Doolittle and Sapienza⁵ (see the previous article) have independently arrived at similar ideas.

The amount of DNA

The large amounts of DNA in the cells of most higher organisms and, in particular, the exceptionally large amounts in certain animal and plant species—the so-called *C* value paradox—has been an unsolved puzzle for a considerable period (see reviews in refs 6–8). As is well known, this DNA consists in part of 'simple' sequences, an extreme example of which is the very large amounts of fairly pure poly d(AT) in certain crabs. Simple sequences, which are situated in chromosomes largely but not entirely in the heterochromatin, are usually not transcribed. Another class of repetitive sequences, the so-called 'intermediate repetitive', have much longer and less regular repeats. Such sequences are interspersed with 'unique' DNA at many places in the chromosome, the precise pattern of interspersions being to some extent different in different species. Leaving aside genes which code for structural RNA of one sort or another (such as transfer RNA and ribosomal RNA), which would be expected to occur in multiple copies (since, unlike protein, their final products are the result of only one stage of magnification, not two), the majority of genes coding for proteins appear to exist in 'single' copies, meaning here one or a few. A typical example would be the genes for α -globin, which occur in one to three copies and the human β -like globins, of which there are four main types, all related to each other but used for slightly different purposes. Notable exceptions are the proteins of the immune system, and probably those of the histocompatibility and related systems. Another exception is the genes for the five major types of histone which also occur in multiple copies. Even allowing for all such special cases, the estimated number of genes in the human genome appears too few to account for the 3×10^9 base pairs found per haploid set of DNA, although it must be admitted that all such arguments are very far from conclusive.

Several authors^{8–13} have suggested that the DNA of higher organisms consists of a minority of sequences with highly specific

functions plus a majority with little or no specificity. Even some of the so-called single-copy DNA may have no specific function. A striking example comes from the study of two rather similar species of *Xenopus*. These can form viable hybrids, although these hybrids are usually sterile. However, detailed molecular hybridization studies show that there has been a large amount of DNA sequence divergence since the evolutionary separation of their forebears. These authors¹³ conclude 'only one interpretation seems reasonable, and that is that the specific sequence of much of the single-copy DNA is not functionally required during the life of the animal. This is not to say that this DNA is functionless, only that its specific sequence is not important'.

There is also evidence to suggest that the majority of DNA sequences in most higher organisms do not code for protein since they do not occur at all in messenger RNA (for reviews see refs 14, 15). Nor is it very plausible that all this extra DNA is needed for gene control, although some portion of it certainly must be.

We also have to account for the vast amount of DNA found in certain species, such as lilies and salamanders, which may amount to as much as 20 times that found in the human genome. It seems totally implausible that the number of radically different genes needed in a salamander is 20 times that in a man. Nor is there evidence to support the idea that salamander genes are mostly present in about 20 fairly similar copies. The conviction has been growing that much of this extra DNA is 'junk', in other words, that it has little specificity and conveys little or no selective advantage to the organism.

Another place where there appears to be more nucleic acid than one might expect is in the primary transcripts of the DNA of higher organisms which are found in the so-called heteronuclear RNA. It has been known for some time that this RNA is typically longer than the messenger RNA molecules found in the corresponding cytoplasm. Heteronuclear RNA contains these messenger RNA sequences but has many other sequences which are never found in the cytoplasm. The phenomenon has been somewhat clarified by the recent discovery of introns in many genes (for a general introduction see ref. 4). Although the evidence is still very preliminary, it certainly suggests that much of the base sequence in the interior of some introns may be junk, in that these sequences drift rapidly in evolution, both in detail and in size. Moreover, the number of introns may differ even in closely related genes, as in the two genes for rat preproinsulin¹⁶. Whether there is junk between genes is unclear but it is noteworthy that the four genes for the human β -like globins, which occur fairly near together in a single stretch of DNA, occupy a region no less than 40 kilobases long¹⁷. This greatly exceeds the total length of the four primary transcripts (that is the four mRNA precursors), an amount estimated to be considerably less than 10 kilobases. There is little evidence to indicate that there are other coding sequences between these genes (although the question is still quite open) and a tenable hypothesis is that much of this interspersed DNA has little specific function.

In summary, then, there is a large amount of evidence which suggests, but does not prove, that much DNA in higher

organisms is little better than junk. We shall assume, for the rest of this article, that this hypothesis is true. We therefore need to explain how such DNA arose in the first place and why it is not speedily eliminated, since, by definition, it contributes little or nothing to the fitness of the organism.

What is selfish DNA?

The theory of natural selection, in its more general formulation, deals with the competition between replicating entities. It shows that, in such a competition, the more efficient replicators increase in number at the expense of their less efficient competitors. After a sufficient time, only the most efficient replicators survive. The idea of selfish DNA is firmly based on this general theory of natural selection, but it deals with selection in an unfamiliar context.

The familiar neo-darwinian theory of natural selection is concerned with the competition between organisms in a population. At the level of molecular genetics it provides an explanation of the spread of 'useful' genes or DNA sequences within a population. Organisms that carry a gene that contributes positively to fitness tend to increase their representation at the expense of organisms lacking that gene. In time, only those organisms that carry the useful gene survive. Natural selection also predicts the spread of a gene or other DNA sequence within a single genome, provided certain conditions are satisfied. If an organism carrying several copies of the sequence is fitter than an organism carrying a single copy, and if mechanisms exist for the multiplication of the relevant sequence, then natural selection must lead to the emergence of a population in which the sequence is represented several times in every genome.

The idea of selfish DNA is different. It is again concerned with the spread of a given DNA within the genome. However, in the case of selfish DNA, the sequence which spreads makes no contribution to the phenotype of the organism, except insofar as it is a slight burden to the cell that contains it. Selfish DNA sequences may be transcribed in some cases and not in others. The spread of selfish DNA sequences within the genome can be compared to the spread of a not-too-harmful parasite within its host.

Mechanisms for DNA spreading

The inheritance of a repeated DNA sequence in a population of eukaryotes clearly requires that the multiplication which produced it occurred in the germ line. Furthermore, any mechanism that can lead to the multiplication of useful DNA will probably lead to the multiplication of selfish DNA (and vice versa). Of course, natural selection subsequently discriminates between multiple sequences of different kinds, but it does not necessarily prevent the multiplication of neutral or harmful sequences.

Multiplication in the germ-line sequence can occur in non-dividing cells or during meiosis and mitosis (within lineages that lead to the germ line). In the former case, the mechanisms available resemble those that are well documented for prokaryotes, that is, multiplication may occur in eukaryotes through the integration of viruses or of elements analogous to transposons and insertion sequences. Doolittle and Sapienza⁵ have discussed these mechanisms in some detail, particularly for prokaryotes. They are likely to lead to the spreading of DNA sequences to widely separated positions on the chromosomes.

During mitosis and meiosis, multiplication (or deletion) is likely to occur by unequal crossing over. This mechanism will often lead to the formation of tandem repeats. It is well documented for the tRNA 'genes' of *Drosophila* and for various other tandemly repeated sequences in higher organisms.

The amount and location of selfish DNA

Natural selection 'within' the genome will favour the indefinite spreading of selfish preferred replicators. Natural selection between genotypes provides a balancing force that attempts to maintain the total amount of selfish DNA at an equilibrium (steady state) level—organisms whose genomes contain an

excessive proportion of selfish DNA would be at a metabolic disadvantage relative to organisms with less selfish DNA, and so would be eliminated by the normal mechanism of natural selection. Excessive spreading of functionless replicators may be considered as a 'cancer' of the genome—the uncontrolled expansion of one segment of the genome would ultimately lead to the extinction of the genotype that permits such expansion. Of course, we do not know whether extinction of genotypes in nature even occurs for this reason.

It is hard to get beyond generalities of this kind. To do so we would, at least, need to know how much selective disadvantage results from the presence of a given amount of useless DNA. Even this minimal information is not easily acquired, so we cannot produce other than qualitative arguments.

It seems certain that the metabolic energy cost of replicating a superfluous short DNA sequence in a genome containing 10^9 base pairs would be very small. If, for example, the selective advantage were equal to the proportion of the genome made up by the extra DNA, a sequence of 1,000 base pairs would produce a selective disadvantage of only 10^{-6} . If the selective disadvantage were proportional to the extra energy cost divided by the total metabolic energy expended per cell per generation, the disadvantage would be much smaller. The selective disadvantage might be greater in more stringent conditions, but it is still hard to believe that a relatively small proportion of selfish DNA could be selected against strongly.

On the other hand, when the total amount of selfish DNA becomes comparable to or greater than that of useful DNA, it seems likely that the selective disadvantage would be significant. We may expect, therefore, that the mechanisms for the formation and deletion of nonspecific DNA will adjust, in each organism, so that the load of DNA is sufficiently small that it can be accommodated without producing a large selective disadvantage. The proportion of nonspecific DNA in any particular organism will thus depend on the lifestyle of the organism, and particularly on its sensitivity to metabolic stress during the most vulnerable part of the life cycle.

We can make one prediction on the basis of energy costs. Selfish DNA will accumulate to a greater extent in non-transcribed regions of the genome than in those that are transcribed. Of course, selfish DNA will in most cases be excluded from translated sequences, because the insertion of amino acids within a protein will almost always have serious consequences, even in diploid organisms (but see the suggestion by F.H.C.C.¹⁸).

At first sight it might seem anomalous that natural selection does not eliminate all selfish DNA. Since the suggestion that much eukaryotic DNA is useless distinguishes the selfish DNA hypothesis from many closely related proposals, it may be useful to take up this point in some detail.

First, the elimination of disadvantaged organisms from a population, by their more favoured competitors, takes a number of generations several times larger than the reciprocal of the selective disadvantage. If the selective disadvantage associated with a stretch of useless DNA in higher organisms is only 10^{-6} , it would take 10^6 – 10^8 years to eliminate it by competition. For typical higher organisms this is a very long time, so the elimination of a particular stretch of selfish DNA may be a very slow process even on a geological time scale. Second, the mechanisms for the deletion of short sequences of DNA may be inefficient, since there is no strong selective pressure for the development of 'corrective' measures when the 'fault' carries a relatively small selective penalty. Taken together, these arguments suggest that the elimination of a particular piece of junk from the genome may be a very slow process.

This in turn suggests that the amount of useless DNA in the genome is a consequence of a dynamic balance. The organism 'attempts' to limit the spread of selfish DNA by controlling the mechanism for gene duplication, but is constrained by imperfections in genetic processes and/or by the need to permit some duplication of advantageous genes. Selfish DNA sequences 'attempt' to subvert these mechanisms and may be able to do so

comparatively rapidly because mutation will affect them directly. On the other hand, the defence mechanisms of the host are likely to depend on the action of protein and therefore may evolve more slowly. Once established within the genome, useless sequences probably have a long 'life expectancy'.

For any particular type of selfish DNA, there is no reason that a steady state should necessarily be reached in evolution. The situation would be continually changing. A particular type of DNA might first spread rather successfully over the chromosomes. The host might then evolve a mechanism which reduced or eliminated further spreading. It might also evolve a method for preferentially deleting it. At the same time, random mutations in the selfish DNA might make it more like ordinary DNA and so, perhaps, less easy to remove. Eventually, these sequences, possibly by now rather remote from those originally introduced, may cease to spread and be slowly eliminated. Meanwhile, other types of selfish DNA may originate, expand and evolve in a similar way.

In short, we may expect a kind of molecular struggle for existence within the DNA of the chromosomes, using the process of natural selection. There is no reason to believe that this is likely to be much simpler or more easy to predict than evolution at any other level. At bottom, the existence of selfish DNA is possible because DNA is a molecule which is replicated very easily and because selfish DNA occurs in an environment in which DNA replication is a necessity. It thus has the opportunity of subverting these essential mechanisms to its own purpose.

The inheritance of selfish DNA

Although the inheritance of selfish DNA will occur mainly within a mendelian framework, it is likely to be different in detail and more complex than simple mendelian inheritance. This is due both to the multiplication mechanisms, which in one way or another will produce repeated copies (see the discussion by Doolittle and Sapienza⁵), and to the fact that these copies are likely to be distributed round the chromosomes rather than being located in a single place in the genome as most normal genes are. For both these reasons, a particular type of selfish DNA is likely to spread more rapidly through a population than would a normal gene with a low selective advantage. It will be even more rapid if selfish DNA can spread horizontally between different individuals in a population, due to viruses or other infectious agents, although it should be remembered that such 'infection' must affect the germ line and not merely the soma. If this initial spread takes place when the additional DNA produced is relatively small in amount, it is unlikely to be seriously hindered by the organism selecting against it. The study of these processes will clearly require a new type of population genetics.

Can selfish DNA acquire a specific function?

It would be surprising if the host organism did not occasionally find some use for particular selfish DNA sequences, especially if there were many different sequences widely distributed over the chromosomes. One obvious use, as repeatedly stressed by Britten and Davidson^{19,20}, would be for control purposes at one level or another. This seems more than plausible.

It has often been argued (see, for example, ref. 21) that for the evolution of complex higher organisms, what is required is not so much the evolution of new proteins as the evolution of new control mechanisms and especially mechanisms which control together sets of genes which previously had been regulated separately. To be useful, a new control sequence on the DNA is likely to be needed in a number of distinct places in the genome. It has rarely been considered how this could be brought about expeditiously by the rather random methods available to natural selection.

A mechanism which scattered, more or less at random, many kinds of repeated sequences in many places in the genome would appear to be rather good for this purpose. Most sets of such sequences would be unlikely to find themselves in the right

combination of places to be useful but, by chance, the members of one particular set might be located so that they could be used to turn on (or turn off) together a set of genes which had never been controlled before in a coordinated way. A next way of doing this would be to use as control sequences not the many identical copies distributed over the genome, but a small subset of these which had mutated away from the master sequence in the same manner.

On this picture, each set of repeated sequences might be 'tested' from time to time in evolution by the production of a control macromolecule (for example, a special protein) to recognize those sequences. If this produced a favourable result, natural selection would confirm and extend the new mechanism. If not, it would be selected against and discarded. Such a process implies that most sets of repeated sequences will never be of use since, on statistical grounds, their members will usually be in unsuitable places.

It thus seems unlikely that all selfish DNA has acquired a special function, especially in those organisms with very high *C* values. Nor do we feel that if one example of a particular sequence acquires a function, all the copies of that sequence will necessarily do so. As selfish DNA is likely to be distributed over the chromosomes in rather a random manner, it seems unlikely that every copy of a potentially useful sequence will be in the right position to function correctly. For example, if a specific sequence within an intron were used to control the act of splicing that intron, a similar sequence in an untranscribed region between genes would obviously not be able to act in this way.

In some circumstances, the sheer bulk of selfish DNA may be used by the organism for its own purpose. That is, the selfish DNA may acquire a nonspecific function which gives the organism a selective advantage. This is the point of view favoured by Cavalier-Smith in a very detailed and suggestive article¹² which the reader should consult. He proposes that excess DNA may be the mechanism the cell uses to slow up development or to make bigger cells. However, we suspect that both slow growth and large cell size could be evolved just as well by other more direct mechanisms. We prefer to think that the organism has tolerated selfish DNA which has arisen because of the latter's own selective pressure.

Thus, some selfish DNA may acquire a useful function and confer a selective advantage on the organism. Using the analogy of parasitism, slightly harmful infestation may ultimately be transformed into a symbiosis. What we would stress is that not all selfish DNA is likely to become useful. Much of it may have no specific function at all. It would be folly in such cases to hunt obsessively for one. To continue our analogy, it is difficult to accept the idea that all human parasites have been selected by human beings for their own advantage.

Life style

The effect of nonspecific DNA on the life style of the organism has been considered by several authors, in particular by Cavalier-Smith¹² and by Hindergardner⁸. We shall not attempt to review all their ideas here but instead will give one example to show the type of argument used.

Bennett²² has brought together the measurements of DNA content for higher herbaceous plants. There is a striking connection between DNA content per cell and the minimum generation time of the plant. In brief, if such an angiosperm has more than 10 pg of DNA per cell, it is unlikely to be an ephemeral (that is, a plant with a short generation time). If it is a diploid and has more than 30 pg of DNA, it is highly likely to be an obligate perennial, rather than an annual or an ephemeral. The converse, however, is not true, there being a fair number of perennials with a DNA content of less than 30 pg and a few with less than 10 pg. A clear picture emerges that if a herbaceous plant has too much DNA it cannot have a short generation time.

This is explained by assuming that the extra DNA needs a bigger nucleus to hold it and that this increases both the size of the cell and the duration of meiosis and generally slows up the development of the plant. An interesting exception is that the

duration of meiosis, is, if anything, shorter for polyploid species than for their diploid ancestors²³. This suggests that it is the ratio of good DNA to junk DNA rather than the total DNA content which influences the duration of meiosis.

An analogous situation may obtain in certain American species of salamander. These often differ considerably in the rapidity of their development and of their life cycles, the tropical species tending to take longer than the more temperate ones. Drs David Wake and Herbert MacGregor (personal communication) tell us that preliminary evidence suggests that species with the longer developmental times often have the higher *C* values. This appears to parallel the situation just described for the herbaceous plants. It remains to be seen if further evidence will continue to support this generalization. (See the interesting paper by Oeldorfe *et al.*²⁵ on 25 species of frogs. They conclude that 'genome size sets a limit beyond which development cannot be accelerated'.)

Testing the theory

The theory of selfish DNA is not so vague that it cannot be tested. We can think of three general ways to do this. In the first place, it is important to know where DNA sequences occur which appear to have little obvious function, whether they are associated with flanking or other sequences of any special sort and how homologous sequences differ in different organisms and in different species, either in sequence or in position on the chromosome. For example, it has recently been shown by Young²⁴ that certain intermediate repetitive sequences in *Drosophila* are often in different chromosomal positions in different strains of the same species.

Second, if the increase of selfish DNA and its movement around the chromosome are not rare events in evolution, it may be feasible to study, in laboratory experiments, the actual molecular mechanisms involved in these processes.

Third, one would hope that a careful study of all the nonspecific effects of extra DNA would give us a better idea of how it affected different aspects of cellular behaviour. In parti-

cular, it is important to discover whether the addition of nonspecific DNA does, in fact, slow down cells metabolically and for what reasons. Such information, together with a careful study of the physiology and life style of related organisms with dissimilar amounts of DNA, should eventually make it possible to explain these differences in a convincing way.

Conclusion

Although it is an old idea that much DNA in higher organisms has no specific function⁸⁻¹², and although it has been suggested before that this nonspecific DNA may rise to levels which are acceptable or even advantageous to an organism^{8,12}, depending on certain features of its life style, we feel that to regard much of this nonspecific DNA as selfish DNA is genuinely different from most earlier proposals. Such a point of view is especially useful in thinking about the dynamic aspects of nonspecific DNA. It directs attention to the mechanisms involved in the spread and evolution of such DNA and it cautions one against looking for a special function for every piece of DNA which drifts rapidly in sequence or in position on the genome.

While proper care should be exercised both in labelling as selfish DNA every piece of DNA whose function is not immediately apparent and in invoking plausible but unproven hypotheses concerning the details of natural selection, the idea seems a useful one to bear in mind when exploring the complexities of the genomes of higher organisms. It could well make sense of many of the puzzles and paradoxes which have arisen over the last 10 or 15 years. The main facts are, at first sight, so odd that only a somewhat unconventional idea is likely to explain them.

We thank W. Ford Doolittle and C. Sapienza for showing us their article before publication, and Drs D. Wake and H. Macgregor for allowing us to quote some of their unpublished conclusions about salamanders. This work was supported by the Eugene and Estelle Ferhauf Foundation, the J. W. Kieckhefer Foundation, the Ahmanson Foundation and the Samuel Roberts Noble Foundation.

1. Dawkins, R. *The Selfish Gene* (Oxford University Press, 1976).
2. Walker, P. M. B. in *Human Genetics: Possibilities and Realities*, 25-38 (Excerpta Medica, Amsterdam, 1979).
3. Crick, F. H. C. in *From Gene to Protein: Information Transfer in Normal and Abnormal Cells* (eds Russell, T. R., Brew, K., Faber, H. & Schultz, J.) 1-13 (Academic, New York, 1979).
4. Crick, F. H. C. *Science* **204**, 264-271 (1979).
5. Doolittle, W. F. & Sapienza, C. *Nature* **284**, 601-603 (1980).
6. Callan, H. G. *J. Cell Sci.* **2**, 1-7 (1967).
7. Thomas, C. A. *Rev. Genet.* **5**, 237-256 (1971).
8. Hinegardner, R. in *Molecular Evolution* (ed. Ayala, F. J.) 179-199 (Sinauer, Sunderland, 1976).
9. Commoner, B. *Nature* **202**, 960-968 (1964).
10. Ohno, S. *J. hum. Evolut.* **1**, 651-662 (1972).
11. Comings, D. E. *Adv. hum. Genet.* **3**, 237-436 (1972).
12. Cavalier-Smith, T. *J. Cell Sci.* **34**, 274-278 (1978).
13. Galan, G. A., Chamberlin, M. E., Hough, B. R., Britten, R. J. & Davidson, E. H. in *Molecular Evolution* (ed. Ayala, F. J.) 200-224 (Sinauer, Sunderland, 1976).
14. Bishop, J. O. *Cell* **2**, 81-86 (1974).
15. Lewin, B. *Cell* **4**, 11-20 (1975); *Cell* **4**, 77-93 (1975).
16. Lomedico, P. *et al. Cell* **18**, 545-558 (1979).
17. Bernards, R., Little, P. F. R., Annison, G., Williamson, R. & Flavell, R. A. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4827-4831 (1979).
18. Crick, F. H. C. *Eur. J. Biochem.* **83**, 1-3 (1978).
19. Britten, R. J. & Davidson, E. H. *Science* **165**, 349-358 (1969).
20. Davidson, E. U. & Britten, R. J. *Science* **204**, 1052-1059 (1979).
21. Wilson, A. C. in *Molecular Evolution* (ed. Ayala, F. J.) 225-236 (Sinauer, Sunderland, 1976).
22. Bennett, M. D. *Proc. R. Soc. B* **181**, 109-135 (1972).
23. Bennett, M. D. & Smith, J. B. *Proc. R. Soc. B* **181**, 81-107 (1972).
24. Young, M. W. *Proc. natn. Acad. Sci. U.S.A.* **76**, 6274-6278 (1979).
25. Oeldorfe, E., Nishioka, M. & Bachmann, K. *Sonderdr. Z.F. Zool., System Evolut.* **16**, 216-24 (1978).

LETTERS

Observation and search for γ rays 1–20 MeV from the Crab, NGC4151, Cyg X-1, Cyg X-3, CG135+1 and 3C273

R. S. White*, B. Dayton†, R. Gibbons*, J. L. Long*,
E. M. Zanrosso* & A. D. Zych*

* Physics Department and Institute of Geophysics and Planetary
Physics, University of California, Riverside, California 92521

† Physics Department, California State University, Los Angeles,
California 90032

Observations of γ rays of 1–20 MeV with the UCR Compton double scatter γ ray telescope are reported for a balloon flight launched from Palestine, Texas, 4.5 GV, at 01.00 UT 29 September 1978. The energy distribution of the total γ rays from the Crab from 1.2 to 20 MeV is measured. Two σ upper limits of 3, 2, 0.6 and 0.4×10^{-4} photons $\text{cm}^{-2} \text{s}^{-1} \text{MeV}^{-1}$ at energies 1.2–3, 3–5, 5–10 and 10–20 MeV, respectively, are found for the Seyfert galaxy NGC4151, the black hole candidate Cyg X-1 and for Cyg X-3. In the same energy intervals, the 2σ upper limits for the nearest and second nearest QSOs CG135+1 and 3C273 are 5, 3, 1 and 0.4×10^{-4} photons $\text{cm}^{-2} \text{s}^{-1} \text{MeV}^{-1}$. These upper limits restrict confirmed γ -ray sources at 1–20 MeV to the Crab and NP0532. Our upper limits for NGC4151 do not support Seyfert galaxies as the source of cosmic diffuse radiation¹.

The University of California UCR Compton double scatter γ -ray telescope measures the energy of the incident γ ray and its scatter angle in the first scintillator. The method of detection, efficiencies and time, angle and energy resolutions have been described previously^{2,3}. Vertical pairs of cells, only, are used in this analysis. Thus, the scatter angle is equal to the zenith angle and the zenith angle of the source is determined for each observed γ ray as a circle on the sky. The sky is divided into circles around the zenith and γ rays are collected into bins in zenith angle and energy. Because of low telescope efficiencies at small angles, $\leq 10^\circ$, and higher backgrounds at large angles $\geq 30^\circ$, we report here only γ rays with zenith angles between 10 and 30° .

Although corrections were made for gain changes in the energy and angle assignments, telescope thresholds are also affected. Therefore, we chose the computer thresholds high enough, 0.3 and 0.9 MeV for S1 and S2, to be above the electronic thresholds throughout most of the flight. Near the beginning of the flight, however, just after reaching the ceiling limit, the computer thresholds were below the actual ones for the small angle scatters and some real γ rays were rejected. The effect is greatest for the 10–20° scatters so we have used no 10–20° data until 30 September 13.00 UT. From 30 September 09.30 UT to flight termination, the residual atmosphere ranged between 4.5 and 3.2 g cm^{-2} . The downward-moving γ -ray background is a combination of atmospheric, cosmic diffuse and neutron-induced backgrounds are nearly independent of height in the atmosphere but we need to correct for the atmospheric γ rays that vary nearly linearly with depth.

The Crab, NGC4151 and 3C273 crossed the zenith meridian at 11.15, 14.75 and 18.15 UT, respectively, 30 September and 1 October; Cyg X-1 and Cyg X-3 transited at 01.45 and 02.30 UT, respectively, 30 September; and CG135+1 transited at 18.00 UT on 1 October as shown in Fig. 1. With few exceptions, observations were made continuously from 09.30 UT on 30 September to 23.00 UT on 1 October.

The count rates for γ rays of 1.2–10 MeV, corrected for altitude variation, were combined into 30-min intervals and plotted in Fig. 1. The data for zenith angles of 10–20° and 20–30° are shown for the duration of the flight at float. As the Earth turns on its axis, a celestial point source passes through the 10–20° and 20–30° telescope zenith angle rings on the sky. The expected shapes of the count rates from a point source are shown by the solid contours.

The background for the Crab has been chosen as an average fit to the data from 30 September, 15.00 UT to 1 October, 07.00 UT and 1 October 16.00 UT to 1 October 23.00 UT. If other sources exist in this region we would be overestimating the background flux and thus underestimating the Crab flux. As our upper limits for other sources is based on this background we would also be overestimating our upper limits for the other sources.

An 8σ increase in count rate is observed in the 10–20° plot and increases of 10σ and 8σ in the 20–30° plot as the Crab transits the zenith meridian.

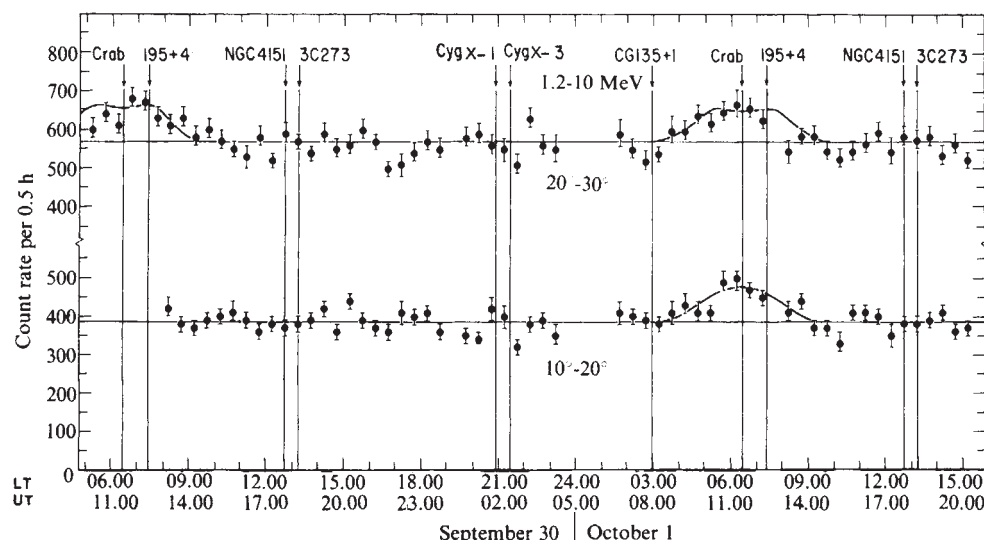


Fig. 1 Counts per 0.5 h versus LT and UT over the flight for γ -ray energies of 1.2–10 MeV and zenith angles of sources of 10–20° and 20–30°. Locations of certain possible sources are indicated by arrows. The contours are the shapes of the γ -ray count rate distributions expected from point sources.

The centre and shape of the combined time distribution of the 10–20° and 20–30° data correlate to the counting rate contour expected from the Crab to within 2 degree of arc, with a 1 σ uncertainty of 3 degree of arc. This eliminates the SAS 2 source, 195+4 and the two COS B sources in the anti-galactic centre direction, CG176–7 and CG189+1, as the origin of the γ rays.

The telescope, with its finite angular resolution, gives a distribution of zenith angles for a point source. The flux in the observation angle rings of 10–20° and 20–30° is calculated at any desired time and the flux missed from outside these rings is estimated and added. For the Crab, at the time of nearest approach, 9°, at the centre of the contours, the flux from 1.2 to 10 MeV is $3.9 \pm 2.0 \times 10^{-3}$ photons cm $^{-2}$ s $^{-1}$.

The energy distribution is determined from plots similar to those in Fig. 1 and plotted in Fig. 2 for the energy intervals of 1.2–3, 3–5, 5–10 and 10–20 MeV. The point at 10–20 MeV gives an upper limit only. Our values are only slightly higher than and compatible with the extrapolations of Walraven *et al.*⁴ and Laros *et al.*⁵ from X-rays to γ -ray energies, as are the recent values of Penningsfeld *et al.*⁶. Our fluxes are about 0.4 of the values of Wilson *et al.*¹¹ measured with the same telescope in 1976 and 0.7 of the preliminary fluxes from this data reported in Kyoto¹. The differences, outside of statistics, arise mostly from the methods of background assignment. Our current values are considered most reliable.

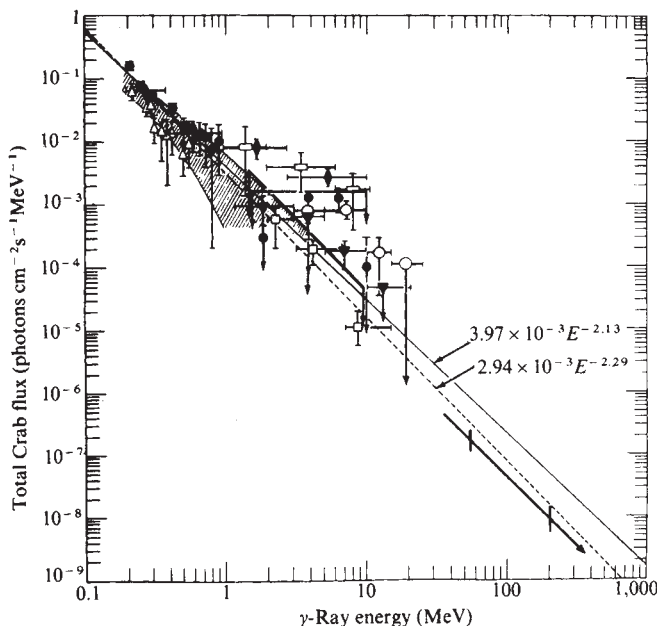


Fig. 2 The energy distribution of X rays and γ rays from the Crab. $\blacktriangledown$, This experiment; $\square$, ref. 6; $\circ$, ref. 11; $\triangle$, ref. 8; $\square$, ref. 12; $\bullet$, ref. 4; $\downarrow$, ref. 9; $\blacklozenge$, ref. 7; $\oplus$, ref. 10; refs 13, 14, $\hline$, fit of Walraven *et al.*⁴; $---$, fit of Laros *et al.*⁵.

NGC4151 crossed within 8° of the zenith at two times separated by 24 h. There is no evidence in our data from the 10–20° or 20–30° angle intervals for an excess flux at either time. For fluxes from NGC4151 2 σ upper limits have been calculated from the background counting rate and are plotted in Fig. 3. These upper limits are factors of 5–10 below the fluxes reported by di Cocco *et al.*¹⁵, Graml *et al.*¹⁶ and Perotti *et al.*¹⁷. The results of Graml *et al.*¹⁶ were withdrawn at the 16th International Cosmic Ray Conference in Kyoto, August 1979, and were replaced by upper limits of 14, 5.5, 2.6 and 0.46×10^{-4} photons cm $^{-2}$ s $^{-1}$ MeV $^{-1}$ at energies of 1.1–2, 2–4, 4–6 and 6–20 MeV, respectively (V. Schönfelder, personal communication). Meegan and Haymes¹⁸ also recently published upper limits for NGC4151. Unless NGC4151 is varying widely in time, by factors of 5–10, our values are in direct conflict with

the above reported fluxes^{15,17} that would have given peaks on Fig. 3 twice as high as the Crab maxima.

As a consequence of our low upper limits and the other upper limits of Fig. 3, the suggestion by Schönfelder²⁶ that the diffuse cosmic γ radiation is the result of superposition of many unresolved galaxies such as the Seyfert galaxy NGC4151 seems premature. There is no observed break in the energy distribution of NGC4151 at 3 MeV and no reason to connect the Seyfert galaxies to the cosmic diffuse γ -ray energy distribution.

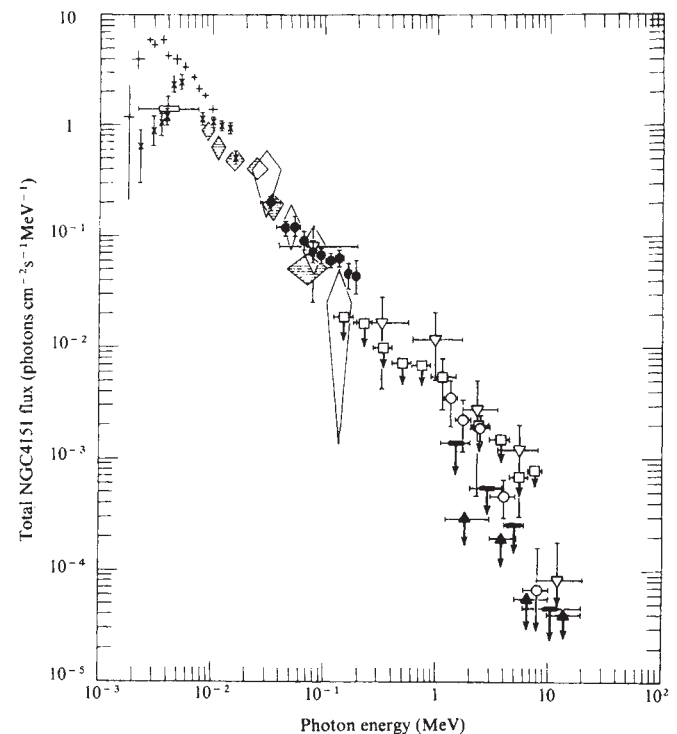


Fig. 3 The energy distribution of X rays and γ rays from NGC4151. $\blacktriangle$, This experiment; $\times$, refs 21, 22; $\square$, ref. 23; $\diamond$, ref. 20; ∇ , ref. 25; $\bullet$, ref. 19; $+$, ref. 24; $\square$, ref. 18; $\circ$, ref. 16; ∇ , ref. 17; $\blacksquare$, F. Graml *et al.* personal communication.

Although Cyg X-1 and Cyg X-3 passed nearly overhead no evidence for either is seen in Fig. 1. The 2 σ upper limits are the same as for NGC4151: 3, 2, 0.6 and 0.4×10^{-4} photons cm $^{-2}$ s $^{-1}$ MeV $^{-1}$ for the energy intervals 1.2–3, 3–5, 5–10 and 10–20 MeV, respectively. For Cyg X-1 our upper limits are a factor of 50 below the fluxes of Baker *et al.*¹² from 1.2 to 10 MeV. Our upper limit from 1.2 to 3 MeV is at the lower limit of the value of Mandrou *et al.*^{13,14}. The fluxes from Cyg X-1 reported by Baker *et al.*¹² would show up on Fig. 1 as peaks 10 times our Crab maxima. It seems that an unusually high time variation, much greater than observed at lower energies (see, for example, ref. 27) would be required to explain the widely varying results of the different observations.

γ -Ray count rates 5 σ above background for energies >150 keV were recently reported from the nearest QSO, CG135+1 (ref. 28); however, no absolute fluxes were given. During our flight CG135+1 transited at a minimum zenith angle of 30.3° and its zenith angle was <35° for 4 h. Somewhat less than half of the count rate is expected in the 10–20° and 20–30° angle intervals. Figure 1 shows no excess γ -ray flux from 1.2 to 10 MeV during this time. Using the same procedure as for the other sources, 2 σ upper limits of 5, 3, 1.0 and 0.4×10^{-4} photons cm $^{-2}$ s $^{-1}$ MeV $^{-1}$ at energies of 1.2–3, 3–5, 5–10 and 10–20 MeV are obtained for CG135+1.

γ -Rays of 50–500 MeV from the second nearest QSO, 3C273, were reported from COS B observations²⁹, but no fluxes

have been measured at energies of 1–20 MeV. During our flight, 3C273 twice transited at a minimum zenith angle of 28.8° and was at zenith angles less than 35° for 3 h each time. No excess is seen in Fig. 1 on either transit. We find the same 2σ upper limits for 3C273 as for CG135 + 1.

We thank NASA for the balloon and helium gas and NASA and NSF for support, also the National Scientific Balloon Facility for the balloon launch, flight and recovery.

Received 31 December 1979; accepted 26 February 1980.

1. Zanrosso, E. M. *et al. Proc. 16th int. Conf. Cosmic Rays*, Kyoto 12, 26–31 (1979).
2. Herzo, D. *et al. Nucl. Instrum. Meth.* **123**, 583–597 (1975).
3. Ryan, J. M. *et al. J. geophys. Res.* **82**, 3593–3601 (1977).
4. Walraven, G. D. *et al. Astrophys. J.* **202**, 502–510 (1975).
5. Laros, J. G., Matteson, J. L. & Pelling, R. M. *Nature phys. Sci.* **246**, 109–111 (1973).
6. Penningsfeld, F. P., Graser, U. & Schönfelder, V. *16th int. Conf. Cosmic Rays*, Kyoto, 1, 101–105 (1979).
7. Gruber, D. E. *Proc. int. Conf. on X-rays in Space* **2**, 875 (1974).
8. Haymes, R. C., Ellis, D. V., Fishman, G. J., Kurfess, J. D. & Tucker, W. H. *Astrophys. J. Lett.* **151**, L9–L14 (1968).
9. Schönfelder, V., Lichti, G. & Moyana, C. *Nature* **257**, 375–376 (1975).
10. Thompson, D. J., Fichtel, C. E., Hartmen, R. C., Kniffen, D. A. & Lamb, R. C. *Astrophys. J.* **213**, 252–262 (1977).
11. Wilson, R. B. *et al. Proc. 15th int. Conf. Cosmic Rays*, Plovdiv 1, 24–29 (1977).
12. Baker, R. E., Lovett, R. R., Orford, K. J. & Ramsden, D. *Nature* **245**, 18–19 (1973).
13. Mandrou, P., Niel, M., Vedrenne, G. & Dupont, A. *Astrophys. J.* **212**, 704–706 (1977).
14. Mandrou, P., Niel, M., Vedrenne, G., Dupont, A. & Hurley, K. *Astrophys. J.* **219**, 288–291 (1978).
15. di Cocco, G. *et al. Nature* **270**, 319–320 (1977).
16. Graml, F., Penningsfeld, P. & Schönfelder, V. *ESA SP-135* 287 (1978).
17. Perotti, F. *et al. Nature* **282**, 484–486 (1979).
18. Meegan, C. A. & Haymes, R. C. *Astrophys. J.* **233**, 510–513 (1979).
19. Auriemma, G. *et al. Astrophys. J. Lett.* **221**, L7–L11 (1978).
20. Baity, W. A., Jones, T. W., Wheaton, W. A. & Peterson, L. E. *Astrophys. J. Lett.* **199**, L5–L8 (1975).
21. Barr, P., White, M. E., Sanford, P. W. & Ives, J. C. *Mon. Not. R. astr. Soc.* **181**, 43 (1977).
22. Elvis, M. *Mon. Not. R. astr. Soc.* **177**, 7–12 (1976).
23. Gursky, H., Kellogg, E. M., Leong, C., Tananbaum, H. & Giacconi, R. *Astrophys. J. Lett.* **165**, L43–L48 (1971).
24. Ives, J. C., Sanford, P. W. & Penston, M. V. *Astrophys. J. Lett.* **207**, L159–L162 (1976).
25. Paciesas, W. S., Mushotsky, R. F. & Pelling, R. M. *Mon. Not. R. astr. Soc.* **178**, 23–25 (1977).
26. Schönfelder, V. *Nature* **274**, 344–346 (1978).
27. Dolan, J. F., Crannel, C. J., Dennis, B. R., Frost, K. J. & Orwig, L. E. *Astrophys. J.* **230**, 551–557 (1979).
28. Barbareschi, L. *et al. Proc. 16th int. Conf. Cosmic Rays*, Kyoto 12, 32–35 (1979).
29. Swanenburg, B. N. *et al. Nature* **275**, 298 (1978).

Multi-ion complexes in the stratosphere—implications for trace gases and aerosol

F. Arnold

Max-Planck-Institut für Kernphysik, 6900 Heidelberg, Postfach 103980, FRG

The fate of stratospheric ions was, until recently, thought to be spontaneous neutralisation following mutual recombination of oppositely charged species. According to Ferguson^{1–3} this need not necessarily apply for sufficiently complex cluster ions, as clustering may stabilise ions against neutralisation. Thus, if neutralisation does not occur, stable ion pairs may be formed. This possibility seemed more likely after the first simultaneous composition measurements of stratospheric positive and negative ions^{4–6} revealed the presence of species more stable than those predicted on theoretical grounds; this led to consideration of a possible role of ion pairs in aerosol formation⁷. Here the formation of ion pairs is investigated in the light of new stratospheric ion composition data. Interactions of ion pairs with other stratospheric constituents, and their significance for aerosol and trace gases is discussed and an experimental check of the existence of the hypothetical stratospheric ion pairs is suggested.

First I consider which ion species among those observed in the stratosphere (positive and negative ions were measured simultaneously around an altitude of 36 km) satisfy the requirements

for stable ion pair formation which is

$$E_D > E_N + E_C \quad (1)$$

Here E_D is the effective energy required to detach the electron from the negative cluster ion, E_N is the effective energy released on neutralisation of the positive cluster ion by a free electron and E_C is the energy released on formation of a chemical bond possibly formed on interaction of core ions or ligands. E_D and E_N are related to the electron affinity E_A of the negative ion core and the ionisation potential IP of the positive ion core by $E_D = E_A + S_-$ and $E_N = IP - S_+$ where S_- and S_+ are the total solvation energies of the negative and the positive cluster ion.

Inspection of the most abundant ions observed around 36 km (Table 1) reveals that the requirement (1) is not fulfilled for the recombination of proton hydrates with $\text{NO}_3(\text{HNO}_3)_2$, as an E_C of 4.5 eV corresponding to an H– HNO_3 bond⁸ which probably forms has to be considered. Recombination of the observed proton hydrates with $\text{HSO}_4(\text{HNO}_3)_2$ probably also leads to spontaneous neutralisation as E_D should not be sufficiently large to compensate for the energy E_C corresponding to an H– HSO_4 bond being 4.5 eV (ref. 9). Thus, recombination of the observed proton hydrate ions does not apparently lead to stable ion pairs. Whether or not the observed ions containing the core HX^+ may form stable ion pairs depends on the proton affinity P_A of the as yet not safely identified molecule X . If, as suggested by Ferguson², X was indeed NaOH due to the extremely high P_A of this molecule the HX^+ ions should form stable ion pairs. If, however, X was another molecule having a considerably lower P_A than NaOH only a fraction of the HX^+ -ion recombination may lead to stable ion pairs, depending on the corresponding E_D and E_N and $\text{PA}(X)$, although exceeding $\text{PA}(\text{H}_2\text{O})$ (ref. 4), may be too low to allow for stable ion pair formation from the observed HX^+ ions.

Note that recombination of more complex cluster ions which have small abundances around 36 km, but which may be even dominant at lower heights should lead to stable pairs even if proton hydrate ions are involved.

In the hypothetical formation of ion pairs, which may be accompanied by loss of some ligand molecules, a stable ion pair interacts with gas molecules. Such interactions may give rise not only to new ligands, but also to a change of core ions. Moreover, due to its large dipole moment the ion pair attaches a free ion quite efficiently. Thus, a three-ion complex may be formed which due to its net charge rapidly attaches an oppositely charged free ion and may form the four-ion complex. Further attachment of free ions may lead to larger multi-ion complexes (MIC) that have the nature of ion crystals or salt particles. An alternative growth mechanism for MIC may be mutual coalescence. Attachment of MIC to pre-existing aerosol particles may also be considered. MIC formation and growth may represent a gas-to-particle formation mechanism which does not require condensation and consequently does not require a super-saturated gaseous species. If, however, a supersaturated species exists, such as sulphuric acid, MIC which have grown to sufficiently large sizes may act as condensation nuclei. The latter have critical sizes which depend on the supersaturation ratio and thus on the abundance of the condensable species and on the temperature. MIC may, therefore, represent a new and interesting class of stratospheric condensation nuclei which had been thought to be mostly of tropospheric and extraterrestrial origin¹⁰. An *in situ* formation of stratospheric condensation nuclei, particularly one initiated by ions was considered to be inefficient^{2,11}. In contrast to the classical mechanism for ion-induced nucleation which requires individual cluster ions to grow to critical sizes and which is inefficient in the stratosphere mainly because there is no gas which is sufficiently super-saturated and is abundant enough to allow for a rapid growth of ions before they recombine. The MIC mechanism avoids this problem by forming condensation nuclei by coalescence of subcritically sized cluster ions. Apart from its possible influence

Table 1 The most abundant positive and negative ions measured around 36 km altitude^{6,16} with their fractional abundances

Positive ion	Abundance (%)	E_N (eV)	Negative ion	Abundance(%)	E_D (eV)
$H^+(H_2O)_4$	21.3	4.03	$NO_3^-(HNO_3)_2$	65.6	5.86
$H^+(H_2O)_5$	17.8	3.38	$HSO_4^-(HNO_3)_2$	14.2	>6.45
$Na^+(H_2O)_4$	19.3	1.92			
$HX^+X(H_2O)_2$	10.7	(<2.53)			
$HX^+X(H_2O)_3$	16.0	(<1.92)			

Values for E_N and E_D are based on measured^{3,17} and estimated cluster ion bond energies. It was assumed that X is NaOH and that NaOH bonds to Na^+ at least as strongly as H_2O and that HNO_3 bonds to HSO_4^- at least as strongly as it bonds to NO_3^- . A measured lower limit to the electron affinity of HSO_4^- (4.5 eV) (ref. 7) was taken. The electron affinity of NO_3^- was taken as 3.9 eV (ref. 18).

on aerosol formation MIC may also have an impact on stratospheric trace gases which may not only condense but also undergo reactions on MIC surfaces. Trace gases which react with, or attach to, these ions may be irreversibly removed from the stratosphere when ions become incorporated into MIC and sediment out of the stratosphere. One example may be sulphuric acid which was found to have a very low gas phase abundance in the stratosphere and which was suspected to be removed by ion processes¹².

If aerosol formation through MIC were significant, it would provide a link between solar activity which controls the stratospheric and tropospheric ionisation sources (galactic and solar cosmic radiation) and stratospheric and possibly also upper tropospheric aerosol. According to Dickinson¹³ such a link seems to be the most favourable way of connecting the Earth's climate to solar activity.

The volume rate R_v at which gas is converted into aerosol by the MIC process is $R_v = \alpha n^2 AM$. Here, α , n and M are the ion-ion recombination coefficient, the total number densities for free positive and negative ions and the average mass of an ion pair, respectively. The parameter A which ranges between 0 and 1 describes the fraction of recombinations which lead to stable ion pairs. Assuming X to be NaOH A becomes 0.5 at 36 km and thus M becomes 4.7×10^{-22} g if α and n are taken as 10^{-7} cm³ s⁻¹ and 2,000 cm⁻³ (ref. 6). For the evaluation of M it was assumed that the mass of an ion pair equals the sum of the masses of the two free cluster ions which mutually recombine. Thus, R_v becomes 1.2×10^{-22} g cm⁻³ s⁻¹. Integration of R_v over the height interval 55–33 km where 'non proton hydrates' have been observed^{4,6} yields a column rate for gas to particle conversion $R_c = 6.3 \times 10^{-17}$ g cm⁻² s⁻¹. Here A was taken to be equal to the fractional abundance of 'non-proton hydrates'⁴ and n_+ and n_- were computed from a mid latitude galactic cosmic ray ionisation rate considering α -values as given by Smith and Church¹⁴. When compared with the column rate for gas-to-particle conversion required to account for the observed stratospheric aerosol layer which is estimated to be about 1.1×10^{-15} g cm⁻² s⁻¹ (ref. 15) the above value is much smaller.

Thus, MIC formation and growth in the upper stratosphere followed by sedimentation of larger MIC into the region of the main aerosol layer apparently cannot account for the total aerosol mass observed in this layer. If, however, at heights below 33 km all ions had the potential to form stable ion pairs ($A = 1$), the MIC process may even account for a large fraction of the total aerosol mass observed.

To estimate MIC abundances it is assumed that ion pairs (2-IC) are formed by mutual recombination and are lost by free ion attachment (rate constant k assumed to be of the order of 10^{-9} cm³ s⁻¹—the upper limit for an ion molecule reaction) to aerosol particles (time constant t ; a sticking coefficient of 1 is assumed) a steady state treatment yields

$$n_2 = \frac{A\alpha n^2}{k2n + t^{-1}} \quad (2)$$

and

$$n_3^+ = A n \left(2 + \frac{1}{knt} \right)^{-1} \quad (3)$$

where n_2 and n_3 are the total number densities of 2-IC and positive 3-IC. Neglecting aerosol interaction for an altitude around 35 km ($A = 0.5$, $n = 2,000$ cm⁻³ and $\alpha = 10^{-7}$ cm³ s⁻¹) n_2 and n_3^+ become 5×10^4 cm⁻³ and 5×10^2 cm⁻³. Analogous steady state considerations lead to $n_4 = 5 \times 10^4$ cm⁻³ and $n_5^+ = 5 \times 10^2$ cm⁻³. A more realistic estimate including aerosol interaction with $t = 10^5$ s (ref. 12) yields $n_2 = 5 \times 10^4$ cm⁻³, $n_3^+ = 1.3 \times 10^2$ cm⁻³, $n_4 = 2.9 \times 10^1$ cm⁻³ and $n_5^+ = 2.9 \times 10^{-1}$ cm⁻³. For negatively charged MIC the number densities should be the same as for positive ones. Thus, for the more realistic case, charged MIC apparently contribute a few per cent to the total concentration of charged particles. Also, the abundance of charged MIC apparently decreases steeply with increasing number of ions contained in these complexes. The estimates, however, depend sensitively on t and thereby on the surface area density of the aerosol, the value of which is somewhat uncertain.

The present detection limit of balloon-borne ion mass spectrometers is ~ 1 particle cm⁻³ for both positively and negatively charged species. When compared with its value, the above estimate of $n_3^+ = n_3^-$ is much larger and, consequently, charged MIC should be detectable even if several individual species of 3-IC are present, and their partial number densities are therefore, considerably smaller than their total number density. Previous instruments were probably not suitable for detection of charged MIC due to their low mass ranges.

Therefore, I suggest a search for stratospheric charged MIC, using ion mass spectrometers with greatly increased mass range. This may provide information relevant to stratospheric aerosol- and trace-gas processes and may be the most promising way of obtaining experimental information about the hypothetical MIC that are difficult to study in the laboratory.

I thank E. Ferguson for helpful discussions.

Received 18 December 1979; accepted 26 February 1980.

1. Castleman, A. W. *Nucleation and Molecular Clustering about Ions* (Report, University of Colorado, Boulder, 1978).
2. Ferguson, E. *Geophys. Res. Lett.* **5**, 1035 (1978).
3. Davidson, J. A., Fehsenfeld, F. C. & Howard, C. J. *Int. J. Chem. Kinet.* **9**, 17 (1977).
4. Arnold, F., Böhringer, H. & Henschen, G. *Geophys. Res. Lett.* **5**, 653 (1978).
5. Arnold, F. & Henschen, G. *Nature* **257**, 521 (1978).
6. Arnold, F., Fabian, R., Henschen, G. & Joos, W. *Planet. Space Sci.* (in the press).
7. Viggiano, A. A., Perry, R. A., Albritton, D. L., Ferguson, E. E. & Fehsenfeld, F. C. *J. geophys. Res.* (in the press).
8. *JANAF, Thermochemical Table* (National Bureau of Standards, PB 168370, Washington DC, 1965).
9. Benson, W. W. *Chem. Rev.* **78**, 23 (1978).
10. Cadle, R. D. & Kiang, C. S. *Rev. geophys. Space Sci.* **15**, 195 (1977).
11. Mohnen, V. A. *Proc. 4th Conf. CIAP (Climatic Impact Assessment Programme)* 1975, 478 (US Department of Transportation, 1975).
12. Arnold, F. & Fabian, R. *Nature* **283**, 55 (1980).
13. Dickinson, R. E. *Bull. Am. Met. Soc.* **56**, 1240 (1975).
14. Smith, D. & Church, M. J. *Planet. Space Sci.* **25**, 433 (1977).
15. Crutzen, P. J. *J. geophys. Res.* **3**, 73 (1976).
16. Arnold, F., Henschen, G., Fabian, R. & Joos, W. *Planet. Space Sci.* (submitted).
17. Kebarle, P. A. *Rev. phys. Chem.* **28**, 445 (1977).
18. Fehsenfeld, F. C., Howard, C. J. & Schmeltekopf, A. L. *J. chem. Phys.* **63**, 2835 (1975).

Jointed blocks of peridotite xenoliths in basalts and mantle dynamics

Asish R. Basu

Department of Geological Sciences, University of Rochester,
Rochester, New York 14627

In considering the formation of fracture in the upper mantle of the Earth, several authors^{1,2} have pointed out that the major problem in the formation of cracks is the fact that tensile stresses in the upper mantle region, particularly in the lithospheric plate, are too small compared with the hydrostatic overburden pressure. From theoretical considerations, Anderson² demonstrated that cracks can nucleate in the lithospheric plates by the accumulation of low density fluid at the base of the plate in the presence of tensile stresses parallel to the plate surface. However, there has been no direct geological observation that supports this conclusion. This letter reports geological observations of planar features in ultramafic inclusions which may have important implications on the existence of fracture in an ambient upper mantle condition.

Ultramafic inclusions in alkalic and related basalts provide much information about the Earth's upper mantle. These inclusions, mostly spinel and garnet peridotites, are considered to be the broken pieces of the wall rocks of the volcanic conduits. In this way, these rocks are the natural 'drill-hole' samples of the upper mantle and may come from depths of up to 200 km (refs 3, 4). Numerous aspects of these rocks have been studied, including the major and minor element chemistry, the isotopic chemistry, the general petrology and petrography, and the rheological properties, to understand the composition and evolution of the upper mantle. However, very little is known about the process by which these inclusions are torn off the walls of the volcanic conduits. The process apparently indicates brittle failure in the mantle in the immediate vicinity of hot magma where the temperatures, in contrast, are appropriate for plastic behaviour.

During petrologic investigation of the ultramafic xenoliths in basalts from San Quintin, Baja California⁵, I observed that many of the xenoliths show angular to subangular shapes resulting from more than one bounding planar surfaces. In some xenoliths three orthogonal sets of such planar surfaces can be recognised. Wilshire and Trask⁶ also recognised similar features in xenoliths from Dish Hill, California. Therefore, planar surfaces in xenoliths in basalts may be common but generally disregarded.

Figure 1 shows a spinel lherzolite xenolith from San Quintin with planar features. The xenolith has a tabular shape, with one perfectly flat surface which measures 25 cm diagonally. Approximately 4 cm away from this surface is another planar feature—a joint plane—that cuts across the massive xenolith. All the bounding surfaces of the xenolith are covered by the frothy, alkalic basalt. As the joint plane does not cut across the host lava, it must have formed before the consolidation of the latter. This observation also suggests that the commonly observed angular to subangular xenoliths result from the formation of sets of joint planes in the upper mantle before invasion by the basalts that carried the xenoliths to the surface. Once the xenoliths are included in this basalt, they are brought up to the surface relatively rapidly by the host basalt as discussed in the following section.

The maximum time spent by the xenoliths in host basalts during the ascent of the magma can be estimated by using Stokes' Law, which states that

$$V_s = \frac{2\Delta\rho ga^2}{9\eta}$$

where V_s is the velocity of settling of a spherical solid of radius a ; $\Delta\rho$, the difference in density between the solid and the fluid; η , the viscosity of the fluid; and g the acceleration due to gravity. Using $a = 15$ cm, the radius of the largest xenolith found in San Quintin, $\Delta\rho = 0.6$ (density of xenolith 3.3 g cm^{-3} and that of basaltic glass $= 2.7 \text{ g cm}^{-3}$) and $\eta = 10^3 \text{ P}$ for alkalic basalt, we obtain $V_s = 30 \text{ cm s}^{-1}$. Therefore, in a static alkali basaltic magma chamber the xenoliths will settle with as high a velocity as 30 cm s^{-1} . For the magmas to ascend to the surface containing the xenoliths, they must ascend with a velocity $> 30 \text{ cm s}^{-1}$. The depths from which the xenoliths come may be as deep as 75 km. Therefore, with a velocity of ascent of 30 cm s^{-1} , the basalt will take 70 h or ~ 3 days to arrive at the surface from a depth of 75 km. As 30 cm s^{-1} is the minimum velocity ascribed to the ascending magma, 3 days is the maximum time the magma takes to arrive at the surface. Therefore, the xenoliths could not have spent more than 3 days in the ascending basalt. An important parameter in the above estimate is the viscosity of alkali basalt taken as 10^3 P . Recent experimental determination⁷ of this parameter suggests that 10^3 P may be higher than the actual viscosity measured under high pressure. In that case, the mean residence time of xenoliths in liquid basalt must be < 3 days. This determination of the relatively short residence time of xenoliths in their host basalts is important because it justifies the interpretation of the chemistry, mineralogy, texture and structure of the xenoliths as representative of an equilibrium condition in the upper mantle.

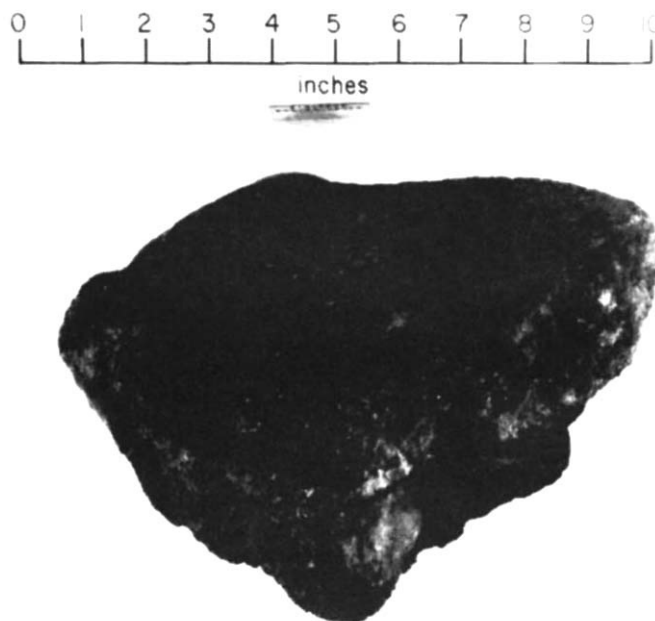


Fig. 1 A jointed block of spinel lherzolite inclusion in alkalic basalt showing a penetrative joint plane parallel to the black marker line. The block of lherzolite is surrounded on all sides by the frothy, scoreaceous basalt that brought up the inclusion to the surface.

There are two possible interpretations for the origin of the joint planes in the xenoliths. First, they could have formed because of the sudden release of lithostatic pressure during the rapid transport to the surface. If that is true, then the joint planes must represent pre-existing planes of weakness. Second, the joints formed by brittle failure in the upper mantle before the xenolith's incorporation by the magma. Note in this connection the mode of occurrence of amphibole selvages on flat surfaces in some xenoliths from Dish Hill, California⁶. The occurrence of amphibole selvages on flat surfaces of these peridotites may indicate that they formed as veins along joint planes in upper mantle conditions.

Is it possible for joint planes to develop under high lithostatic pressure in the upper mantle? It is well known that the presence

of fluid within a massive body of rock reduces the strength of the rock^{8,9}. In effect, the fluid pressure reduces the load pressure, thereby reducing the strength of the rock⁸. This principle has been applied to the effective decrease in strength of a rock due to the generation of magma and to the process of crack propagation by magma in the brittle failure region¹¹. The field of brittle fracture is expanded by fluid pressure because the pressure of the fluid in pores and cracks counteracts the normal stress across the cracks from lithostatic pressure. However, there is no apparent evidence for the presence of fluid magma or volatiles along the joint plane of the lherzolite xenolith under discussion. In the absence of the evidence of such fluid phases along the joint plane, the alternative and most viable explanation for the origin of the joint planes is that they formed under high differential stress of the order of several kilobars.

Analysis of stresses in ultramafic xenoliths showing varying degrees of plastic deformation indicate a range of 250 to 700 bar and analyses of upper mantle strain rates show that rates of 10^{-15} to 10^{-13} s^{-1} are typical for plastically deformed mantle as represented by ultramafic xenoliths¹². However, the joint planes in the xenolith under discussion must have formed at much higher stresses and the corresponding strain rates must be significantly higher than 10^{-13} s^{-1} .

An argument can be made against the interpretation that the joint planes in the xenolith formed due to the release of lithostatic pressure. During the transport of the xenoliths to the surface from mantle depths, two competing processes will operate: decompression due to pressure release and the contraction due to fall in temperature. The anisotropy of olivine, the major constituent of the xenoliths, is of the order of 5% (ref. 13). Thus, the low differential compressibility of olivine in the xenolith is expected to be manifest in the scale of the individual grains. The net effect of these processes can be observed in the mineral grains of many xenoliths which become uncompact due, possibly, to this differential compressibility. The scale of the observed joint planes (several centimetres) in the xenolith (Fig. 1) indicate then the presence of directed stress of at least a similar scale.

Therefore, it is possible that the joint plane (Fig. 1) in the lherzolite xenolith formed in the upper mantle by brittle failure; some intermediate focus earthquakes such as those in Hawaii¹⁴ may indicate this brittle behaviour. When the alkalic magma rises from below, it picks up the jointed blocks of the mantle and brings them to the surface relatively rapidly. In this way, the jointed blocks of the xenoliths will still preserve the ambient properties of the upper mantle, such as high P , T conditions of mineralogical equilibria and evidence of plastic flow.

The most probable explanation for the origin of the joint plane in the lherzolite (in Fig. 1) is that it formed because of high differential stress in the upper mantle. The commonly observed angular to subangular shapes of mantle-derived xenoliths in alkali basalt reflect this phenomenon, and provide a mechanism for these mantle rocks to be torn off the walls of the volcanic conduits.

This research was supported by NSF grant (EAR 76-26569). I thank Peter Hudleston and Ian MacGregor for helpful discussions.

Received 4 October 1979; accepted 26 February 1980.

- Wertman, J. J. *geophys. Res.* **76**, 8544 (1971).
- Anderson, O. L. *Kimberlites, Diatremes and Diamonds: Their Geology, Petrology, and Geochemistry*, 344 (American Geophysical Union, 1979).
- Boyd, F. R. *Geochim. cosmochim. Acta* **37**, 2533 (1973).
- MacGregor, I. D. & Basu, A. R. *Science* **185**, 1007 (1974).
- Basu, A. R. thesis, Univ. California (1975).
- Wilshire, H. A. & Trask, N. J. *Am. Miner.* **56**, 240 (1971).
- Kushiro, I., Yoder, H. S. & Mysen, B. O. *Trans. Am. geophys. Un.* **57**, 354 (1977).
- Hubbert, M. K. & Rubey, W. W. *Bull. geol. Soc. Am.* **70**, 115 (1959).
- Raleigh, C. B. & Paterson, M. S. *J. geophys. Res.* **70**, 3965 (1965).
- Brace, W. F. *Am. Geophys. Un. Mon.* **16**, 265 (1972).
- Yoder, H. S. Jr, *Generation of basaltic magma*, 265 (National Academy of Sciences, Washington, 1976).
- Poisy, P. & Nicholas, A. *Nature* **274**, 429 (1978).
- Carter, N. L., Baker, D. W. & George, R. P. *Am. Geophys. Un. Monogr.* **16**, 167 (1972).
- Eaton, J. P. & Murata, K. J. *Science* **132**, 925 (1960).

Uppermost Miocene carbon isotope stratigraphy of a piston core in the equatorial Pacific

L. D. Keigwin Jr

Graduate School of Oceanography, University of Rhode Island, Kingston, Rhode Island 02881

N. J. Shackleton

Godwin Laboratory, Freeschool Lane, Cambridge, UK

Carbon isotopic analyses of benthonic foraminifera from palaeomagnetically dated piston core RC12-66 (east equatorial Pacific Ocean) show a permanent decrease in $\delta^{13}\text{C}$ of $\sim 0.5\%$ in the upper reversed interval of palaeomagnetic Epoch 6 (about 6.3 Myr BP). Magnetostratigraphic and biostratigraphic correlations reported here suggest that this decrease is isochronous with a similar carbon shift observed at localities elsewhere in the Indo-Pacific region.

A permanent decrease of ~ 0.5 – 0.8% in the $\delta^{13}\text{C}$ of benthonic foraminiferal calcite has been reported from the Pacific¹⁻⁴ and Indian Oceans⁵ during the late Miocene. These studies are part of the CENOP (Cenozoic Palaeoceanography) programme which is seeking to define isochronous markers for palaeoceanographic reconstructions⁶. This shift in $\delta^{13}\text{C}$ has been interpreted as an oceanwide change in the metabolic CO_2 content of deep waters, resulting from a change in circulation and upwelling patterns, from increased supply of organic matter to the deep sea due to a late Miocene regression, and from a change in ocean fertility^{2,5,7}. Palaeomagnetic dating of this event at an uplifted shallow-marine sequence (Blind River) in New

Table 1 Carbon isotopic composition of *G. subglobosa*, core RC12-66

Depth (cm)	$\delta^{13}\text{C}$ (% PDP)	Depth (cm)	$\delta^{13}\text{C}$ (% PDB)
1,730	-0.57	2,195	-0.95
1,870	-0.79	2,237	-0.60
1,875	-1.00	2,299	-0.56
1,877.5	-0.91	2,319	-0.91
1,887.5	-1.03	2,352	-0.82
1,890	-0.76	2,359	-0.62
1,915	-0.73	2,380	-1.08
1,962.5	-0.46	2,399	-0.89
1,975	-1.01	2,419	-0.66
1,985	-0.43	2,439	-1.02
2,015	-0.36	2,445	-0.98
2,027	-0.55	2,458	-0.63
2,050	-0.79	2,462	-0.30
2,062.5	-0.73	2,478	-0.64
2,077.5	-0.71	2,499	-0.70
2,085	-0.61	2,519	-0.30
2,092.5	-0.65	2,535	-0.78
2,102.5	-0.57	2,539	-0.23
2,125	-0.97	2,559	-0.59
2,135	-0.51	2,560	-0.70
2,147	-0.55	2,579	-0.73
2,155	-0.98	2,598	-0.13
2,165	-0.91	2,602	-0.38
2,175	-0.72	2,627	-0.31
2,185	-0.55	2,690	-0.21
		2,782	-0.22

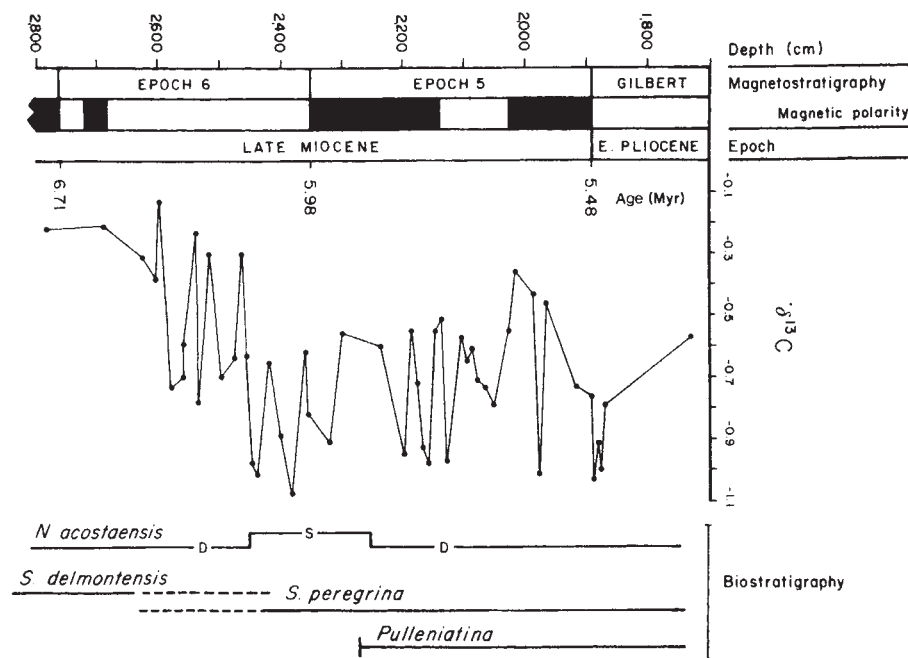


Fig. 1 Carbon isotopic results (‰, PDB), magnetostratigraphy¹⁰, and biostratigraphy¹¹ of the latest Miocene at core RC12-66.

Zealand places it at ~6.2 Myr, within palaeomagnetic Epoch 6 (refs. 3, 8, 9). It has still not been proved that the late Miocene carbon shift represents an isochronous event because it has been palaeomagnetically dated only at the Blind River section. From biostratigraphic evidence, however, the late Miocene carbon shift was assigned ages of ~6.5 Myr (ref. 1) and ~6.2 Myr (ref. 5).

We have found the carbon shift in east Equatorial Pacific Ocean piston core RC12-66. Core RC12-66 (02°36.6' N, 148°12.8' W, 4755m) is one of the longest piston cores recovered from the deep ocean (28 m). This core has been the subject of detailed magnetostratigraphic¹⁰ and biostratigraphic studies^{11,12}. Magnetic plug samples (of ~5 cm³) originally studied by Foster and Opdyke¹⁰ were provided by L. H. Burckle of Lamont-Doherty Geological Observatory. The samples were washed and specimens of the benthonic foraminifer *Globocassidulina subglobosa* were picked from the size fraction greater than 175 µm and analysed using standard procedures¹³.

Carbon isotopic results (Table 1), expressed as $\delta^{13}\text{C}$ with respect to the standard PDB, decrease upwards by about 0.5‰ between ~2,450 and 2,600 cm (Fig. 1). There is unexplained high frequency variability of as much as 0.6‰, but on average the shift seems permanent. Pliocene and Pleistocene $\delta^{13}\text{C}$ results from nearby piston core V28-179 (ref. 14) are on average lighter than the lowest results at RC12-66. Likewise, other sections in the Indo-Pacific show that this is indeed a permanent change.

The age that is applied to the 1.5-m interval over which the shift occurs depends on the age assignment of the palaeomagnetic boundaries of magnetic Epoch 6. Both this study and that of Loutit and Kennett use a new time scale from the Gilbert/Gauss boundary to the inferred base of Epoch 6 based on direct dating of subaerially erupted lavas in Iceland¹⁵. Ages for the boundaries of Epoch 6 given by McDougall *et al.*¹⁵ and corrected in accordance with Mankinen and Dalrymple¹⁶ are ~5.98 and 6.71 Myr. Using this time scale, which is probably the most accurate for magnetic Epoch 6, we estimate the age of the late Miocene carbon shift to be ~6.3 Myr. The corrected time scale of McDougall *et al.*¹⁵ indicates an age of ~6.2 Myr for the carbon shift in the Blind River section³, in good agreement with that for RC12-66. Magnetostratigraphy, therefore, suggests that the carbon shift is isochronous between the south-west Pacific land-based sequence and the east equatorial Pacific Ocean.

Biostratigraphy also suggests the carbon shift is an isochronous marker in deep-sea sequences. At RC12-66 the carbon shift interval is marked by the evolution of the radiolarian *Stichocorys peregrina* from *Stichocorys delmontensis*¹¹. Just above this level there is an interval of sinistral *Neoglobobulimina acostaensis*, followed by the first appearance of *Pulleniatina primalis*. This is the same sequence of events found at DSDP 158 (ref. 1). In addition, where nanofossil biostratigraphy has been examined, it is found that the carbon shift is associated with the first appearance of *Ammalolithus primus*¹⁷.

At DSDP Site 158 the carbon shift was originally reported to occur within a 3-m interval¹, but subsequent more detailed studies by one of us (L.D.K.) reveal that it is restricted to only a 1-m interval. The sedimentation rate through this interval is estimated to be ~50m Myr⁻¹ (ref. 18), and thus the carbon shift seems to have occurred in ~50,000 yr. In the New Zealand Blind River section an upper limit on the duration of the shift is 150,000 yr. (ref. 3). At RC12-66 we estimate from the magnetostratigraphy of Foster and Opdyke¹⁰, that the late Miocene sedimentation rate is ~8m Myr⁻¹. This rate provides a duration of this event of ~200,000 yr at RC12-66. Further studies of high sedimentation rate cores are necessary to establish the minimum duration of this event.

We thank J. P. Kennett and T. S. Loutit and E. Vincent for comments. L. D. K. was supported by NSF grant OCE 76-81489 (CENOP).

Received 28 December 1979; accepted 27 February 1980.

1. Keigwin, L. D. Jr *Earth planet. Sci. Lett.* **45**, 361-382 (1979).
2. Bender, J. L. & Keigwin, L. D. Jr *Earth planet. Sci. Lett.* **45**, 383-393 (1979).
3. Loutit, T. S. & Kennett, J. P. *Science* **204**, 1196-1199 (1979).
4. Loutit, T. S. *Mar. Micropaleont.* (in the press).
5. Vincent, E., Killingley, J. S. & Berger, W. H. *Mar. Micropaleont.* (in the press).
6. Savin, S. M. & Sommer, M. A. *GSA A. Meet. Abstr. Prog.* **11**, 510 (1979).
7. Bender, M. L. & Keigwin, L. D. Jr *GSA A. Meet. Abstr. Prog.* **11**, 387 (1979).
8. Kennett, J. P. & Watkins, N. D. *Bull. geol. Soc. Am.* **85**, 1385-1398 (1974).
9. Ryan, W. B. F. *et al. Riv. Ital. Paleont.* **80**, 631-688 (1974).
10. Foster, J. H. & Opdyke, N. D. *J. geophys. Res.* **75**, 4465-4473 (1970).
11. Saito, T., Burckle, L. H. & Hays, J. D. in *Late Neogene Epoch Boundaries* (eds Saito, T. & Burckle, L. H.) 226-244 (American Museum of Natural History, New York, 1975).
12. Gartner, S. *Bull. geol. Soc. Am.* **84**, 2021-2034 (1973).
13. Shackleton, N. J. *CNRS Colloq.* **219**, 203-209 (1974).
14. Shackleton, N. J. & Opdyke, N. D. *Nature* **270**, 216-219 (1977).
15. McDougall, I. *et al. Bull. geol. Soc. Am.* **88**, 1-15 (1977).
16. Mankinen, E. A. & Dalrymple, G. B. *J. geophys. Res.* **84**, 61-626 (1979).
17. Haq, B. *et al.* (in preparation).
18. Heath, G. R. & van Andel, T. J. *Init. Rep. DSDP* **16**, 899-913 (1973).

North Atlantic oceanography as possible cause of Antarctic glaciation and eutrophication

Detmar Schnitker

Department of Oceanography, University of Maine at Orono, Walpole, Maine 04573

Tarling¹ suggested that ice ages may occur when continental blocks occupy polar positions, isolating the pole from oceanic influences, and when nearby seas are available to provide moisture for the build-up and maintenance of polar ice caps. Such conditions gradually became established in the Southern Hemisphere where the present ice age began 13 Myr ago². The Cenozoic circum-Antarctic geographic and oceanographic evolution has been recently summarised by Kennett³. I report here that circum-Antarctic palaeogeography and palaeoceanography alone were insufficient to lead to the Antarctic glaciation. A model is presented that suggests that the final triggering events took place in the North Atlantic.

Formation of the Atlantic and Indian oceans during the Cretaceous and the early Cenozoic left Antarctica and Australia in a polar position⁴. Separation of Australia from Antarctica during the late Eocene opened a seaway in which an almost continuous circum-Antarctic current system could form, which in turn effectively isolated the Antarctic continent⁵. Concomitant with, and perhaps caused by this change of ocean configuration was a decrease of global warmth from a mid-Cretaceous high⁶, so that by late Eocene/earliest Oligocene time freezing temperatures prevailed along Antarctic coasts⁷. Although the southern ocean has probably been the principal source of the world oceans' deep and bottom water since at least the late Cretaceous⁸, the formation of Antarctic bottom water at freezing temperatures near the Eocene/Oligocene boundary stimulated the oceans' deep water circulation as manifested by widespread erosional hiatuses⁹, a drastic drop of the carbonate compensation depth^{10,11}, and a reorganisation of benthic deep water faunas^{12,13}. During the remainder of the Palaeogene Antarctica remained a cold but largely ice-free continent.

Zonal winds (polar easterlies, prevailing westerlies), driven by the newly established thermal contrast, set up zonal ocean currents. After opening of the Drake passage near the Oligocene/Miocene boundary, ~22.5 Myr ago¹⁴, a fully circum-Antarctic current formed. The Ekman drift that resulted from these zonal wind fields caused upwelling around Antarctica. This upwelling, however, could have brought up only cold water of low salinity from shallow (~100 m) depth (A. L. Gordon, personal communication). A probable consequence of this upwelling and of the interaction of the vigorous currents with bottom and coastal features is a zone of high biological productivity, noted to have occurred from the late Eocene onward, in proximity to the Antarctic continent¹⁵⁻¹⁷. But in general, the southern ocean remained an oligotrophic region of general downwelling, dominated by calcareous floras and faunas. The great transition to essentially modern conditions, that is Antarctic continental glaciation and circum-Antarctic oceanic eutrophication, occurred during the early and middle Miocene. These events have no apparent 'cause' in any Southern Hemisphere tectonic or oceanographic event, but seem to correlate well with two tectonic events in the Northern Hemisphere that fundamentally rearranged the North Atlantic circulation system: (1) the closure of the Mediterranean at its eastern end, and (2) the subsidence of the Iceland ridge (Greenland-Iceland-Faeroe ridge).

Closure of the eastern Mediterranean during the early Miocene was completed ~18 Myr ago¹⁸, interrupting the circum-

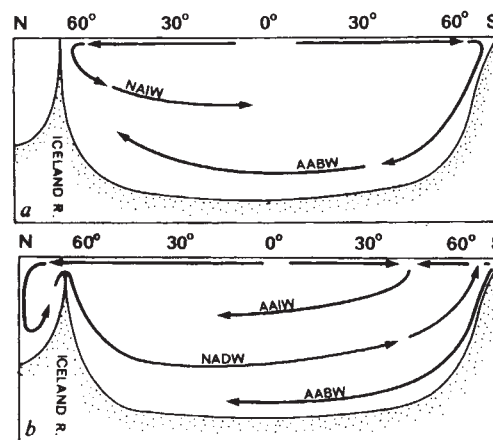


Fig. 1 General current system of the Atlantic Ocean. *a*, Pre-Miocene, with emergent Iceland ridge. *b*, Post-mid-Miocene, with submerged Iceland ridge and Norwegian Sea overflow.

global, near-equatorial flow through Tethys. The resulting sea, lying at or below 30° N latitude⁴, became an evaporation basin which must have contributed increasing amounts of warm and very saline water to the North Atlantic⁸. During the early Miocene the Atlantic probably became, as it still is today, the saltiest of all oceans.

Since seafloor spreading between Greenland and Europe began, ~53 Myr ago, the aseismic Iceland ridge had been an effective barrier between the Atlantic ocean and the north polar sea. During the latest Oligocene, ~25 Myr ago, the first segments of the Iceland ridge began to subside below sea level^{19,20}. This initial subsidence allowed ingress of North Atlantic surface water into the then already cold Norwegian sea²¹, but the barrier was probably still too shallow to allow return flow of deep water into the North Atlantic. By the mid-Miocene, subsidence of the Iceland ridge had progressed to such depths that reflux of cold deep water became significant, thus linking the north polar sea as a heat sink to the world ocean. The establishment of the North Atlantic as a major source of deep water for the world ocean brought about a fundamental rearrangement of the global deep-water circulation.

The early Norwegian Sea overflow water, formed through cooling of North Atlantic water, might not have been very cold, having achieved sufficient density for sinking and outflow as a result of its high salinity which it inherited from the Mediterranean outflow. The outflow of this early Norwegian Sea overflow water soon reached great proportions, causing major sediment redistribution within the North and South Atlantic ocean²²⁻²⁴.

The early North Atlantic deep water which resulted from the Norwegian Sea overflow water, like its modern counterpart, travelled the length of the Atlantic Ocean and inserted itself from below as an intermediate water mass into the circum-Antarctic system. The arrival of this North Atlantic deep water destabilised the vertical structure of the circum-Antarctic system so that the existing wind-induced belt of upwelling could bring North Atlantic deep water to the surface instead of the cold, low salinity local water from shallow depths (A. L. Gordon, personal communication). The upwelling water was nutrient-rich and led to the transformation of an oligotrophic area to one of high biological productivity, which has characterised Antarctic waters until today.

The great volume of relatively warm but saline North Atlantic deep water that was thus injected into the cold circum-Antarctic environment contained heat that could be converted to latent heat by evaporation. The ensuing high evaporation rates supplied much moisture to the Antarctic continent, which had been 'precooled' to below freezing temperatures for nearly 25 Myr, since the late Eocene. Retention of this moisture on Antarctica formed and maintained the continental ice cap. This

sudden build-up of the Antarctic ice cap as soon as sufficient moisture became available suggests that the previous lack of substantial glaciation on Antarctica was probably due to severe dryness rather than moderate temperatures.

Increased thermal contrast in the Antarctic induced stronger wind fields which led to an enlargement of the upwelling system, as indicated by the progressive northward displacement of the Antarctic convergence³ and also to continued growth of the Antarctic ice sheet, which by the end of the Miocene may have been 50% larger than today²⁵. The eustatic lowering of sea level which accompanied the growth of the Antarctic ice sheet was sufficient to bring about the final separation of the Mediterranean from the Atlantic, the Messinian salinity crisis²⁶. The Mediterranean, therefore, ceased to supply saline water to the North Atlantic. North Atlantic surface water now flowing into the Norwegian-north polar sea required, because of its lessened

salinity, more intensive cooling to achieve sufficient density for sinking and later overflow. The eustatically lowered sea level also shoaled the Iceland ridge and probably caused a reduction of the volume of outflow of this new, colder, and less saline Norwegian Sea overflow water. When this new North Atlantic deep water was brought to the surface by the circum-Antarctic upwelling system it maintained, because of its high nutrient content, the very productive Antarctic siliceous bioprovince, but because of its lower temperature reduced the rates of evaporation from the ocean. Precipitation 'starvation' led to a drastic reduction of the Antarctic ice sheet soon after the Miocene-Pliocene boundary^{27,28}.

I thank Dr A. L. Gordon for explaining important aspects of Antarctic physical oceanography, also Drs J. P. Kennett and T. B. Kellogg for critical reviews, and the NSF for support through grant DES 75-13855.

Received 26 September 1979; accepted 21 February 1980.

1. Tarling, D. H. in *Climatic Change* (ed. Gribbin, J.) 3-24 (Cambridge University Press, 1978).
2. Denton, G. H., Armstrong, R. L. & Stuiver, M. in *Late Cenozoic Glacial Ages* (ed. Turekian, K. K.) 267-306 (New Haven, 1971).
3. Kennett, J. P. *J. geophys. Res.* **82**, 3843-3860 (1977); *Mar. Micropal.* **3**, 301-345 (1978).
4. Smith, A. G. & Briden, J. C. *Mesozoic and Cenozoic Paleogeographic Maps* (Cambridge University Press, 1977).
5. Kennett, J. P. *et al. Init. Rep. DSDP* **29**, 1155-1169 (1975).
6. Savin, S. M. A. *Rev. Earth planet. Sci.* **5**, 319-355 (1977).
7. Kennett, J. P. & Shackleton, N. J. *Nature* **260**, 513-515 (1976).
8. Schnitker, D. *Earth Sci. Rev.* (in the press).
9. Moore, T. C., van Andel, T. H., Sancetta, C. & Pisias, N. *Micropaleontology* **24**, 113-138 (1978).
10. van Andel, T. H. *Earth planet. Sci. Lett.* **26**, 187-194 (1975).
11. van Andel, T. H., Thiede, J., Sclater, J. G. & Hay, W. W. *J. Geol.* **85**, 651-698 (1977).
12. Benson, R. H. *Lethaia* **8**, 69-83 (1975).

13. Schnitker, D. *Init. Rep. DSDP* **48**, 377-413 (1979).
14. Barker, P. F. & Burrell, J. *Mar. Geol.* **75**, 15-34 (1977).
15. Hayes, D. E. & Frakes, L. A. *Init. Rep. DSDP* **28**, 919-942 (1975).
16. Tucholke, B. F. *et al. Init. Rep. DSDP* **35**, 359-400 (1976).
17. Burns, D. A. *Mar. Geol.* **25**, 35-59 (1977).
18. Berggren, W. A. & Hollister, E. D. in *Studies in Paleooceanography* (ed. Hay, W. W.) 126-186 (SEPM, Special Publication 20, 1974).
19. Vogt, P. R. *Nature* **239**, 79-81 (1972).
20. Talwani, M. & Udintsev, G. *Init. Rep. DSDP* **38**, 1213-1242 (1976).
21. Schrader, H. J., Björklund, K., Manum, S., Martini, E. & van Hinte, J. *Init. Rep. DSDP* **38**, 1197-1211 (1976).
22. Jones, E. J. W., Ewing, M., Ewing, J. I. & Eittrheim, S. L. *J. geophys. Res.* **75**, 1655-1680 (1970).
23. Ruddiman, W. F. *Bull. geol. Soc. Am.* **83**, 2039-2062 (1972).
24. Tucholke, B. E. & Ewing, M. *Bull. geol. Soc. Am.* **85**, 1789-1802 (1974).
25. Ryan, W. B. F., Cita, M. B., Rawson, M. D., Burckle, L. H. & Saito, T. *Riv. Ital. Paleont.* **80**, 631-688 (1974).
26. Shackleton, N. J. & Kennett, J. P. *Init. Rep. DSDP* **29**, 801-807 (1975).
27. Mayewski, P. A. *Rep. 56* (Ohio State University, Institute of Polar Studies, 1975).

***In vitro* effects of high molecular weight forms of ACTH on the fetal sheep adrenal**

Margaret M. Roebuck & Colin T. Jones

Nuffield Institute for Medical Research, Headley Way, Oxford OX3 9DS, UK

Diane Holland & Robert Silman

Department of Reproductive Physiology, The Medical College of St Bartholomew's Hospital, London EC1, UK

The direct involvement of the pituitary-adrenal axis in birth has been well established, at least in sheep, and its removal prolongs pregnancy^{1,2}. As part of the process the fetal sheep adrenal grows rapidly during the 10-15 d prepartum and is associated with a large rise in the plasma corticosteroid concentration^{3,4}. This does not seem to result from an increased ACTH secretion^{5,6}. The fetal adrenal *in vivo* seems refractory to circulating ACTH and shows poor response to elevation of plasma concentration^{1,2,7-9}. Thus the signal for the adrenal hypertrophy and the initiation of parturition remains unclear. The responsiveness of the fetal adrenal to ACTH has been re-examined using isolated adrenal cells. The study shows that in the fetal sheep these are not inherently unresponsive to ACTH, but that high-molecular-weight forms of ACTH block the action of ACTH₁₋₃₉. These peptides may be responsible for controlling the activity of the adrenal *in situ*.

To see whether the pathway for ACTH stimulation of steroidogenesis was poorly developed in the fetal sheep adrenal 10-30 d before birth cells were prepared from the gland and incubated *in vitro*. Cells from fetal sheep of 120-135 d were prepared by collagenase digestion^{10,11}. Adrenals were trimmed, weighed, finely chopped with a scalpel blade and suspended in

10 ml of Krebs bicarbonate buffer containing 20 mM-HEPES pH 7.4, 0.5% (w/v) bovine serum albumin and 25 mg of collagenase (Sigma, Type 1) in a 20-ml plastic vial. This was gassed with 95% O₂ and 5% CO₂ at 37 °C for 60 min and the cell suspension stirred with a polytetrafluoroethylene paddle rotating at 500 r.p.m. The suspension was filtered through a 100-μm nylon mesh and centrifuged at 200g for 15 min. The cells were washed twice in 10 ml of Krebs bicarbonate buffer containing 0.5 mg ml⁻¹ soya bean trypsin inhibitor (Sigma), then finally suspended in 50 ml of the same buffer containing the trypsin inhibitor. Aliquots of 0.5 ml of this cell suspension were added to 0.1 ml of 50 mM sodium phosphate (pH 7.5) containing a range of ACTH concentrations in 2-ml polystyrene tubes, then incubated under 95% O₂ and 5% CO₂ for 2 h at 37 °C in a shaking water bath. At the end of the incubation the cells were centrifuged at 1,000g for 15 min at 4 °C and the supernatant was used for the assay of cortisol. Details of the collection of plasma samples and of the assay of ACTH and cortisol have been described previously⁷. The ACTH₁₋₃₉ used in these studies was the human standard obtained from the National Institute for Biological Standards and Control, London.

For the preparation of large molecular weight peptides, pituitaries were collected from adult Welsh mountain sheep. They were rapidly frozen in liquid nitrogen and stored at -70 °C. Before chromatography they were thawed and homogenised in 0.02 M-HCl (10 mg in 0.5 ml) with 1 mM disopropylphosphorofluoridate. After centrifugation at 2,500 r.p.m. for 10 min, 0.5-2 ml of the supernatant was run on a 1.6 × 100 cm column of Sephadex G-100 Superfine and eluted with 1 M acetic acid at 1.6 ml per h and 4 °C. Fractions were collected every 15 min. Three large molecular weight peptides A, B and C, which react with anti-ACTH₁₋₂₄, were separated¹². They have approximate molecular weights of 50,000, 30,000 and 20,000 respectively¹². In some cell incubations they were added either in place of ACTH₁₋₃₉ or together with 166 pg ml⁻¹ of ACTH₁₋₃₉.

Previous experiments on the fetal sheep with chronically implanted vascular catheters have shown relatively small changes in corticosteroid concentration for relatively large rises

Table 1 The output of cortisol from adrenal cells of 120–135-d fetal sheep in response to various ACTH concentrations

	0	ACTH concentration (pg ml ⁻¹)			
		40	166	332	1,660
Cortisol produced (ng per 2 h per g of DNA)	0.4±0.16	1.45±0.4	4.35±1.35	5.86±1.64	6.73±1.98
Percentage increase over baseline	—	222%	867%	1,202%	1,396%

The results are means ±s.e.m. of nine experiments.

in ACTH⁷⁻⁹. In contrast to the *in vivo* data, adrenal cells prepared from fetal sheep of 120–135 d gestation respond by causing a marked increase in cortisol output in response to low concentrations of ACTH₁₋₃₉ (Table 1). Note particularly the effect of increasing the concentration from 40 to 330 pg ml⁻¹ and to 1,660 pg ml⁻¹ which represents approximately the basal to the elevated concentration of ACTH₁₋₃₉ in plasma of fetal sheep for experiments carried out *in vivo*⁷⁻⁹. An interpretation of the *in vivo* compared with the *in vitro* data is that *in vivo* something is suppressing the ability of the fetal adrenal to respond to circulating ACTH, although a direct comparison is not easy. This does not appear to be a low PO₂ (unpublished observations). The proportional increase in cortisol output by the fetal adrenal cells between 40 and 330 pg of ACTH per ml,

was of the same order as that found in adrenal cells prepared from the glands of adult guinea pigs, man, rats and sheep¹³⁻¹⁵.

The pituitary of the fetal sheep contains ACTH-like peptides of high molecular weight which are secreted^{12,16} and found in the circulation¹⁷. The proportion of these to ACTH₁₋₃₉ falls in the preparturient period¹⁸. The effects of these peptides on cortisol responses of adrenal cells have shown consistent results in four fetal sheep (Fig. 1). They illustrate that the large peptides (even A, *P*<0.05) have steroidogenic activity at 1 ng ml⁻¹, but inhibit the response of the ACTH₁₋₃₉ (Fig. 1). It is difficult to exclude the possibility that there is proteolytic breakdown of the large peptides to ACTH₁₋₃₉, as occurs in plasma; were that occurring *in vitro*, it would mask part of the inhibitory action of the three peptides, thus accounting for the impression of the absence of inhibition in some cases.

These three peptides are present in the pituitary of the fetal sheep and are secreted by pituitary cells. The concentration of high-molecular-weight ACTH-like peptides in the plasma (mostly peptides similar to A and B) of fetal sheep at 120–135 d is about 150 pg ml⁻¹ compared with a value of about 50 pg ml⁻¹ for ACTH₁₋₃₉ (ref. 18).

Thus the existence of these peptides in the fetal circulation may explain why the fetal adrenal is apparently less responsive *in vivo* than *in vitro*. Moreover, the change in the ratio of the concentration of these to that of ACTH₁₋₃₉ that occurs late in gestation in the fetal sheep could be involved in the initiation of parturition, possibly by influencing the trophic and stereogenic effects of ACTH₁₋₃₉ on the adrenal. If these big peptides are suppressing the fetal adrenal of the sheep then they provide an

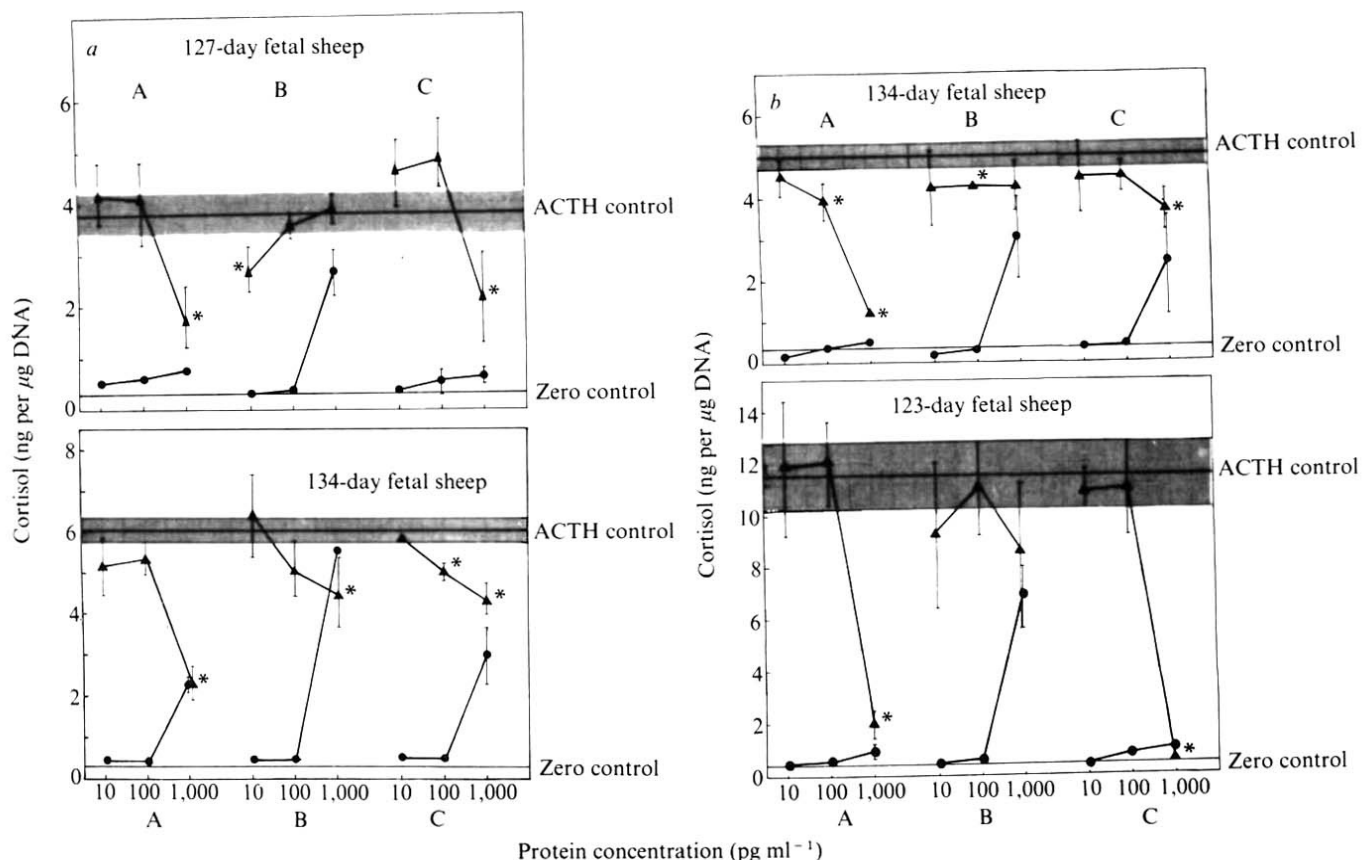


Fig. 1 The response, to ACTH₁₋₃₉ and various ACTH peptides, of adrenal cells prepared from four fetal sheep of 123–134 d gestation. Cells were incubated for 2 h with either 166 pg ml⁻¹ ACTH₁₋₃₉ (ACTH control; the shaded area is equivalent to 2 s.d.), with ACTH₁₋₃₉ plus various concentrations of big peptides (▲), with the big peptide alone (●), or without any addition (zero concentration control). The peptides A, B and C have molecular weights of about 50,000, 30,000 and 20,000, respectively. The values are means ±s.d. for three separate incubations. The rates of cortisol production in the zero controls were 123 d 0.43±0.03, 127 d 0.35±0.07, 134 d 0.39±0.02, 134 d 0.28±0.02 ng of cortisol per 2 h per µg DNA.

* Indicates a significant difference between the response of the cells incubated with ACTH₁₋₃₉ at 100 pg ml⁻¹ as compared with that on incubation with the appropriate peptide plus 100 pg ml⁻¹ of ACTH₁₋₃₉ (*P*<0.05).

explanation for the paradoxical effect of short-term dexamethasone infusion into the fetal sheep increasing the response to ACTH₁₋₂₄ (ref. 19) as dexamethasone would reduce the secretion of the large peptides as well as that of ACTH₁₋₃₉.

The mechanism by which these large peptides suppress adrenal steroidogenic response to ACTH is not clear. Provided they contain the 15–18 region of the ACTH molecule they should compete effectively for the ACTH receptor and then may block the accessibility of ACTH₁₋₃₉ to the steroidogenic site on the adrenal cell membrane²⁰.

We are grateful to Professor G. S. Dawes for his interest and encouragement. M.M.R. is an MRC Scholar. This study was supported by grants from the MRC and the Wellcome Trust.

Received 19 November 1979; accepted 15 February 1980.

- Liggins, G. C., Fairclough, R. J., Grieves, S. A., Kendall, J. Z. & Knox, B. S. *Rec. Prog. Horm. Res.* **29**, 111–150 (1973).
- Bassett, J. M. & Thorburn, G. D. in *The Endocrinology of Pregnancy and Parturition. Experimental Studies in Sheep* (ed. Pierrepont, C. G.) 126–140 (Alpha Omega Alpha Publishing, Cardiff, 1973).
- Bassett, J. M. & Thorburn, G. D. *J. Endocr.* **44**, 285–286 (1969).
- Comline, R. S., Nathanielsz, P. W., R. B. & Silver, M. J. *Physiol., Lond.* **210**, 141P–142P (1970).
- Rees, L. H., Jack, P. M. B., Thomas, A. L. & Nathanielsz, P. W. *Nature* **253**, 274–275 (1975).
- Jones, C. T., Boddy, K. & Robinson, J. S. *J. Endocr.* **72**, 293–300 (1977).
- Jones, C. T., Boddy, K., Robinson, J. S. & Ratcliffe, J. G. *J. Endocr.* **72**, 279–292 (1977).
- Jones, C. T., Robinson, R. O., Luther, E., Ritchie, J. W. K. & Worthington, D. J. *Endocr.* **73**, 11–20 (1977).
- Jones, C. T. & Rurak, D. J. *Endocr.* **70**, 327–328 (1976).
- Sayers, G., Swallow, R. L. & Giordano, N. D. *Endocrinology* **88**, 1063–1068 (1971).
- Lowry, P. J., McMartin, C. & Peters, J. J. *Endocr.* **59**, 43–55 (1973).
- Silman, R. E. *et al. J. Endocr.* **81**, 19–34 (1979).
- Kolanowski, J., Ortega, N. & Crabbe, J. *Annls Endocr.* **35**, 501–507 (1974).
- Kolanowski, J. & Crabbe, J. *Molec. cell Endocr.* **5**, 255–267 (1976).
- Beale, R. J. & Sayers, G. *Archs Biochem. Biophys.* **148**, 70–76 (1972).
- Jones, C. T. & Ritchie, J. W. K. *J. Endocr.* (submitted).
- Jones, C. T. *J. Endocr.* **71**, 38–39 (1976).
- Jones, C. T. & Roebuck, M. J. *Steroid Biochem.* (in the press).
- Liggins, G. C., Fairclough, R. J., Grieves, S. A., Foster, C. S. & Knox, B. S. *CIBA Fdn Symp.* **47**, 5–25 (1977).
- Hoffman, K. *Am. Handbook Physiol.* 7 sect. Part II, 29–58 (1974).

Recipients were young adult males, whose aorta and vena cava in mid-abdomen was slit anteriorly and joined by sutures of 9.0 nylon to the cuffs of the donor aorta and vena cava. The urine from the added kidney was drained into the recipient's bladder by sewing the patch of donor bladder, which included the ureteral orifices, into a defect cut into the dome of the bladder with 7.0 Dexon. Thus there was no surgical interference with the blood vessels or ureters of the recipient's own kidneys. Unless the donor kidney is spared from ischaemia during transplantation, the resulting acute tubular necrosis of the transplant is particularly severe and the transplanted kidneys atrophy³. We perfused donor kidneys with ice-cold Ringer solution and performed rapid anastomoses. We discarded rats in which the transplants atrophied, as well as those that developed hydro-nephrosis or bladder stones.

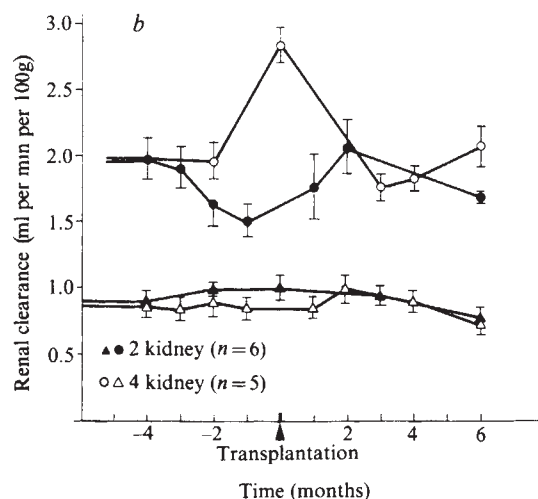


Fig. 1 *a*, Excretory urogram of a four-kidney rat 6 months after transplantation (25 min after injection of iodinated diatrizoate). The added (below) and original kidneys have maintained their size and excrete similar concentrations of contrast. *b*, Sequential renal clearances of ¹²⁵I-iothalamate (GFR, triangles) and ¹³¹I-Hippuran (RPF, circles) relative to body weight in five four-kidney rats (open symbols) and six intact sibling controls (closed symbols). Urograms were prepared with standard diagnostic X-ray equipment on rats anaesthetised with ether, 15–20 min after intravenous injection into the penile vein of 1.5 ml of Renographin 76 (67% diatrizoate, 37% bound iodine, Squibb).

Autoregulation in rats with transplanted supernumerary kidneys

Ruben F. Gittes, Marcus Rist, Salvador Treves & Andrew Biewiner

Harvard Program in Urology (Longwood Area), Department of Surgery, Peter Bent Brigham Hospital, and Departments of Radiology, Children's Hospital Medical Center, Harvard Medical School, Boston, Massachusetts 02115

Autoregulation of mammalian renal function is implicit in the fact that a few hours after removal of one kidney the activity of the intact nephrons in the other kidney suddenly increases, both in glomerular filtration and tubular reabsorption. There follows a slower but easily measured increase in renal mass—compensatory hypertrophy in each residual nephron. That response to a reduction of nephrons has been well studied¹, but the adaptation of normal renal function to additional nephrons could not be tested until microsurgical techniques made it possible to transplant one or two additional kidneys into normal rats². We report here sequential measurements of the total renal function of rats with three and four kidneys. Our results indicate that autoregulation maintains total renal function at normal levels in spite of a doubling in the number of nephrons and apparent renal mass.

Strictly inbred Lewis rats were used to avoid immunological reaction against the added tissue and to facilitate long-term observations. As before², the donor kidney (or kidneys) was removed *en bloc* with adjacent segments of aorta and vena cava, as well as with the entire ureter and a small patch of bladder.

The success of transplantation and the cross-sectional area of the kidneys were demonstrated on excretory urograms during the first week, at 4–6 months and at 1 yr (Fig. 1a). Cross-sectional areas and measurements made with calipers at transplantation and autopsy, showed that in four-kidney rats there was no obvious shrinkage of the original kidneys. We conclude that normal nephron size and aggregate renal mass do not adjust downwards as a consequence of a doubling of the number of nephrons.

Six four-kidney rats were used for long-term studies of glomerular filtration rate (GFR) and plasma renal flow (PRF). We added two kidneys because the doubling of functional mass should override the variability common to the results of clearance studies. A femoral or neck vein was cannulated with polyethylene tubing (PE-50) under ether anaesthesia and the rats were placed in individual metabolic cages (Nuclear Corp.). Saline was infused overnight (2 ml h^{-1} Harvard Instrument infusion pump). The next day, if the urinary output of the awake but restricted rats matched their input, saline was replaced by an isotonic solution of ^{125}I -sodium iothalamate (Glofil, Abbott) and ^{131}I -sodium iodohippurate (^{131}I -Hippuran, NEN) to measure GFR and RPF, respectively. Infusion was continued for another 5 h. Urine produced during the first hour was discarded and then all urine was collected for 4 h: its volume was measured and its radioactivity counted. To obtain blood counts we took blood at the beginning and the end of the collection period by warming the tails and making a cut in each of them. The two isotopes were measured simultaneously in a γ counter (Nuclear Chicago) with appropriate settings. The clearance of both isotopes was calculated as ml min^{-1} and adjusted for body weight. Although simultaneous counting of both isotopes introduced a systematic apparent lowering of the absolute values of RPF, it did not affect the comparison with the controls similarly processed.

GFR and RPF were measured three times in 4 months preoperatively and again at monthly intervals after surgery for 4 more months, and at 6 months for surviving rats. Total GFR and RPF increased with age and with increasing body weight during those 10 months in both groups, but not as a consequence of adding two kidneys. The stability of their GFR and RPF relative to body weight before and after transplantation is shown in Fig. 1b.

We next measured the relative contribution of each kidney towards overall clearance of the isotopes from the blood. We used dimercaptosuccinic acid (DMSA) a mercurial compound taken up selectively by the proximal convoluted tubules in the renal cortex. When bonded to the γ -emitter technetium-99 ($^{99\text{m}}\text{Tc}$ -DMSA), it is used clinically for scintiscanning of functioning renal cortex, usually to compare the uptake of two kidneys in one individual. We examined the uptake of $^{99\text{m}}\text{Tc}$ -DMSA in each kidney of six rats with three-kidneys, 7–9 months after transplantation, and of their intact sibling controls. We injected approximately $1 \mu\text{Ci}$ of $^{99\text{m}}\text{Tc}$ -DMSA (Medi Physics) intravenously: 16 h later anterior and posterior images (Fig. 2a) were recorded by a γ camera with a pinhole collimator (Sigma 400, Ohio Nuclear) and stored on a digital computer (Gamma-11, Digital Equipment). The count rate of the syringe containing the isotope was also stored on the computer before and after injection to determine the activity injected. Renal uptake was expressed for each kidney as a percentage of activity injected, corrected for background. Each value was an average of the uptake measured on the anterior and posterior projections.

Total uptake of DMSA varied from 21 to 30.5% of the injected dose, and in paired controls from 23.7–35%, thus being the same in rats with three kidneys as in a normal rat (Fig. 2b). Uptake by each of the two original kidneys varied from 7.6 to 12.3% of the injected dose, while for the controls the range was 10.3–18.2%. Paired measurements, made on the same day, always showed the uptake of each kidney in the control rat to be greater than that of either of the two original kidneys in the rat with three kidneys. The group difference on analysis by a

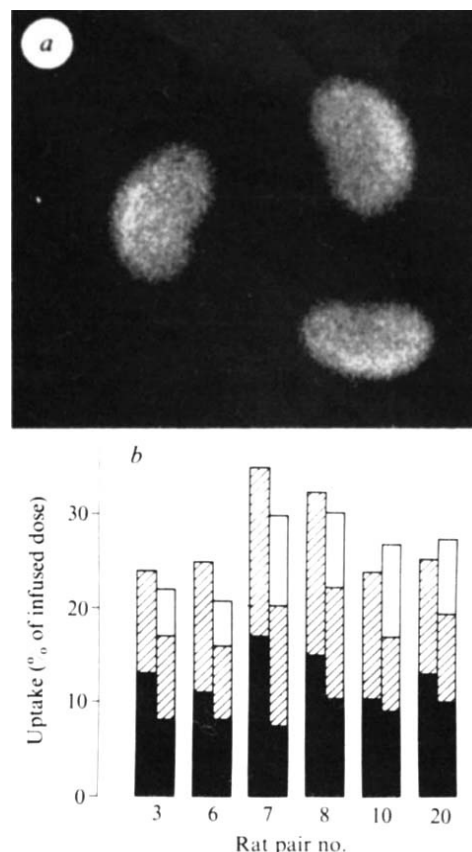


Fig. 2 a, Scintigram of three-kidney rat 16 h after injection of $^{99\text{m}}\text{Tc}$ -DMSA, in posterior view. The right kidney is uppermost and the transplanted third kidney lies transversely below the other two. b, Uptake of DMSA by the six three-kidney rats and sibling pair controls. The total uptakes of the three-kidney rats (right bars) are the same as those of controls (left bars). But the uptakes of the original two kidneys (the right kidney is black and the left is hatched in each bar) are consistently reduced ($P < 0.001$) in the three-kidney rats to accommodate the function of the added third kidney (clear sections of right-hand bars).

binomial probability table⁴ for a sample of 12 with no overlap is highly significant ($P < 0.0002$).

Taken together, our measurements demonstrate that the individual function of intact kidneys was reduced in response to the addition of functional renal tissue. Silber and Crudup tentatively concluded that a third kidney increased GFR by 50% (ref. 5), and an increase was suggested but not proved in our earlier work³. Our findings now conclusively show no increase in rats with three or four kidneys. Kidneys can increase their glomerular filtration and tubular reabsorption work acutely without an increase in size, as noted after uninephrectomy⁶. It now seems that on a day-to-day basis the kidney is capable of a wide range of autoregulation of function, including downward modulation, in response to the excretory needs of the whole body.

A feedback signal seems to enable the aggregate renal glomerular and tubular clearance of rats with two, three or four kidneys to remain a constant linear function of body weight. The signal is probably humoral because the transplanted kidneys responded as well as the intact kidneys. Another hint of a feedback signal comes from the work of Kover and Tost⁷, who connected the circulation of dogs to an isolated two-kidney organ preparation, acutely reducing the urine volume and urine sodium excretion of the original kidneys in proportion to that of the isolated kidneys.

The lack of detectable decrease in kidney size is consistent with the concept that a minimum renal size is achieved during normal growth and is 'obligatory'⁸, not influenced by low functional demand. Thus a baby (4 weeks) rat with a transplanted adult kidney grows up and has normal sized kidneys

(unpublished results of R. A. Dieppa, R.F.G., A.B. and J. McDonald). By contrast, compensatory hypertrophy is reversible to the 'obligatory' baseline size: transplanting a normal kidney into a rat living on a greatly hypertrophied half of one kidney (after removal of one and a half kidneys) reverses the hypertrophy of the half kidney⁹.

This work was supported in part by Thornburg Medical Research Foundation, the Brigham Surgical Group Foundation, Inc., Schweizerische Akademie der medizinischen Wissenschaften (MR) and the Division of Nuclear Medicine, Children's Hospital Medical Center. We thank Dr Sun Lee, Sandy Koch, Jeanne McDonald, Yvette Landre and Audrey Samul for technical assistance.

Received 10 October 1979; accepted 26 February 1980.

1. Kaufman, J. M., Siegel, N., Lytton, B. & Hayslett, J. P. *Invest. Urol.* **13**, 441–444 (1976).
2. Lee, S., Gittes, R. F. & Orloff, M. J. *Am. J. Surg.* **127**, 296–299 (1974).
3. Klein, T. W. & Gittes, R. F. *J. Urol.* **109**, 19–27 (1973).
4. Freund, J. E. *Mathematical Statistics* 2nd edn, 426 Prentice-Hall, Englewood Cliffs, 1971).
5. Silber, S. & Crudup, J. *Invest. Urol.* **11**, 466–470 (1974).
6. Weinman, E. J., Renquist, K., Stroup, R., Kashgarian, M. & Hayslett, J. P. *Am. J. Physiol.* **224**, 565–571 (1973).
7. Kover, G. & Tost, H. *Acta physiol. acad. sci. hung.* **5**, 41–50 (1978).
8. Silber, S. & Malvin, R. L. *Am. J. Physiol.* **226**, 114–117 (1974).
9. Mitchell, T. S., Lee, S. & Gittes, R. F. *Invest. Urol.* **12**, 213–217 (1974).

Dopaminergic nerve endings visualised by high-resolution autoradiography in adult rat neostriatum

Laurent Descarries, Olivier Bosler, France Berthelet & Michel H. Des Rosiers

Centre de Recherche en Sciences Neurologiques and Département de Physiologie, Faculté de Médecine, Université de Montréal, Montréal, Québec, Canada H3C 3T8

Nerve cells containing dopamine (DA), noradrenaline (NA) or serotonin possess powerful physiological mechanisms for the reuptake and storage of their own transmitter. Monoaminergic neurones may therefore be visualised by high-resolution autoradiography¹, provided that tritiated biogenic amines are retained in their sites of uptake and/or storage during suitable preparative histological procedures. Using this method, noradrenergic and serotonergic neuronal cell bodies and axon terminals have been specifically identified and examined in various parts of mammalian central nervous system². However, although the nerve cell bodies of central dopaminergic neurones have been similarly investigated after *in vivo* labelling with either ³H-DA or ³H-NA^{3–7}, the autoradiographic detection of dopaminergic nerve endings has remained problematic. Indeed, in previous studies after cerebroventricular infusion, superfusion or local instillation of ³H-DA or ³H-NA, light and electron microscope autoradiographs of neostriatum, the brain region richest in dopaminergic nerve endings, failed to exhibit the small clusters of silver grains typical of monoamine-labelled axonal varicosities^{8–10}. This suggested tracer displacement during histological processing, so we tested, after intraventricular administration of ³H-DA, a technique of tissue preparation developed in our laboratory for high-resolution autoradiographic visualisation of deoxyglucose-6-phosphate, a diffusible, water-soluble, unbound substrate¹¹. We report here that the rapid successive vascular perfusion of both a glutaraldehyde primary fixative and osmium tetroxide postfixative, which may be followed by dehydration and resin-embedding also carried out by perfusion, maintains a sufficient amount of ³H-DA *in situ* to ensure light and electron microscopic identification of DA axon terminals in rat brain¹².

All experiments were carried out in adult Sprague-Dawley rats (200–400 g), pretreated with a monoamine oxidase inhibitor (β -phenylisopropylhydrazine, 5 mg per kg intraperitoneally, 18 and 2 h earlier) and anaesthetised with sodium pentobarbital, 100–200 μ Ci of 3,4-dihydroxy [*ring*-G-³H]phenylethylamine

hydrochloride (³H-DA, specific activity 6.3 Ci mmol⁻¹, Amersham), diluted in 100–200 μ l of saline containing 1% ascorbic acid, were instilled stereotactically into a lateral cerebral ventricle for 1–2 h. In some cases, non-radioactive NA or serotonin (10⁻³ M) was added to this solution at a molarity 2–10 times higher than that of the tracer. Immediately thereafter, fixation was initiated by high flow-rate aortic-arch perfusion with 600 ml of 3.5% glutaraldehyde in 0.1 M cacodylate buffer (200 ml min⁻¹) and continued without interruption by perfusion with 500 ml of 0.5% osmium tetroxide in 0.2 M cacodylate buffer (same rate). Subsequent dehydration and embedding were also carried out by perfusion, using 3–4 l absolute ethanol and 3 $\times$ 200 ml Epon-propylene oxide mixtures at increasing resin concentrations. The whole brain was then immersed for several hours in pure Epon and polymerised *in toto* at 60 °C¹¹. Samples of the caudate nucleus, ipsilateral to the instilled ventricle, were processed as semi-thin (1 μ m-thick) and thin (silver-gold) sections for light or electron microscope autoradiography, respectively, according to standard dipping techniques¹³. Other brain regions were examined by light microscope autoradiography only.

Three types of autoradiographic reactions were detected in neostriatum after intraventricular administration of ³H-DA. (1) Diffuse labelling, in the form of dispersed silver grains, overlaid all cytological constituents of the paraventricular region, with relative sparing of myelinated bundles. This ubiquitous reaction spread over 150–750 μ m and attenuated gradually with increasing distance from the ependyma. Its intensity was markedly diminished if non-radioactive NA had been added to the ³H-DA solution, indicating a common affinity of both catecholamines for nonspecific and/or artefactual binding sites. (2) Ependymal labelling (not illustrated) consisted of clumps of silver grains seemingly aligned along the ventricular edge in light microscope autoradiographs. In electron microscope autoradiographs, silver grains in this location mostly overlaid intracellular lysosome-like corpuscles predominating in the apical portion of ependymocytes, and plurivesicular or dense granular blebs protruding at the ventricular surface. Such labelling was not found after the addition of either non-radioactive NA or serotonin to the ³H-DA solution, suggesting competition between all three biogenic amines for binding at these ependymal sites. (3) Small and dense aggregates of silver grains, disseminated throughout the neuropil of paraventricular neostriatum, were the main finding of the present study (Fig. 1a). These were typical of labelled axonal varicosities having taken up and retained tritiated biogenic amine in high concentration¹⁴. They were most numerous immediately beneath the ependyma and also densely grouped in deeper patch-like areas extending between some myelinated fibre bundles. Electron microscope autoradiography confirmed that all such accumulations of tracer involved axonal enlargements containing synaptic vesicles. In fact, after a short period of exposure and low-sensitivity development, these were practically the only sites of silver grain localisation (Fig. 1b).

The above experiments give some indication of the signal value of the axonal accumulations of ³H-DA. First, these were no longer visible when non-radioactive NA had been added to the ³H-DA solution, even though rat neostriatum receives little if any NA innervation¹⁵. This finding was consistent with the notion of competitive uptake of the two catecholamines in dopaminergic nerve endings^{16–18}. Second, the addition of non-radioactive serotonin to the ³H-DA solution had no apparent effect on either the number or the labelling intensity of the axonal accumulations, as would have been expected from a cross-reactivity of serotonergic nerve endings, abundant in neostriatum. Accordingly, there was no apparent labelling, even after the administration of ³H-DA alone, neither of the supraependymal serotonergic plexus nor of serotonergic axonal varicosities in supra-chiasmatic nucleus, two regions readily accessible to tracer. The latter observations were at variance with previous *in vitro* data indicating that serotonergic fibres have a capacity to take up exogenous DA¹⁹. The apparent discrepancy

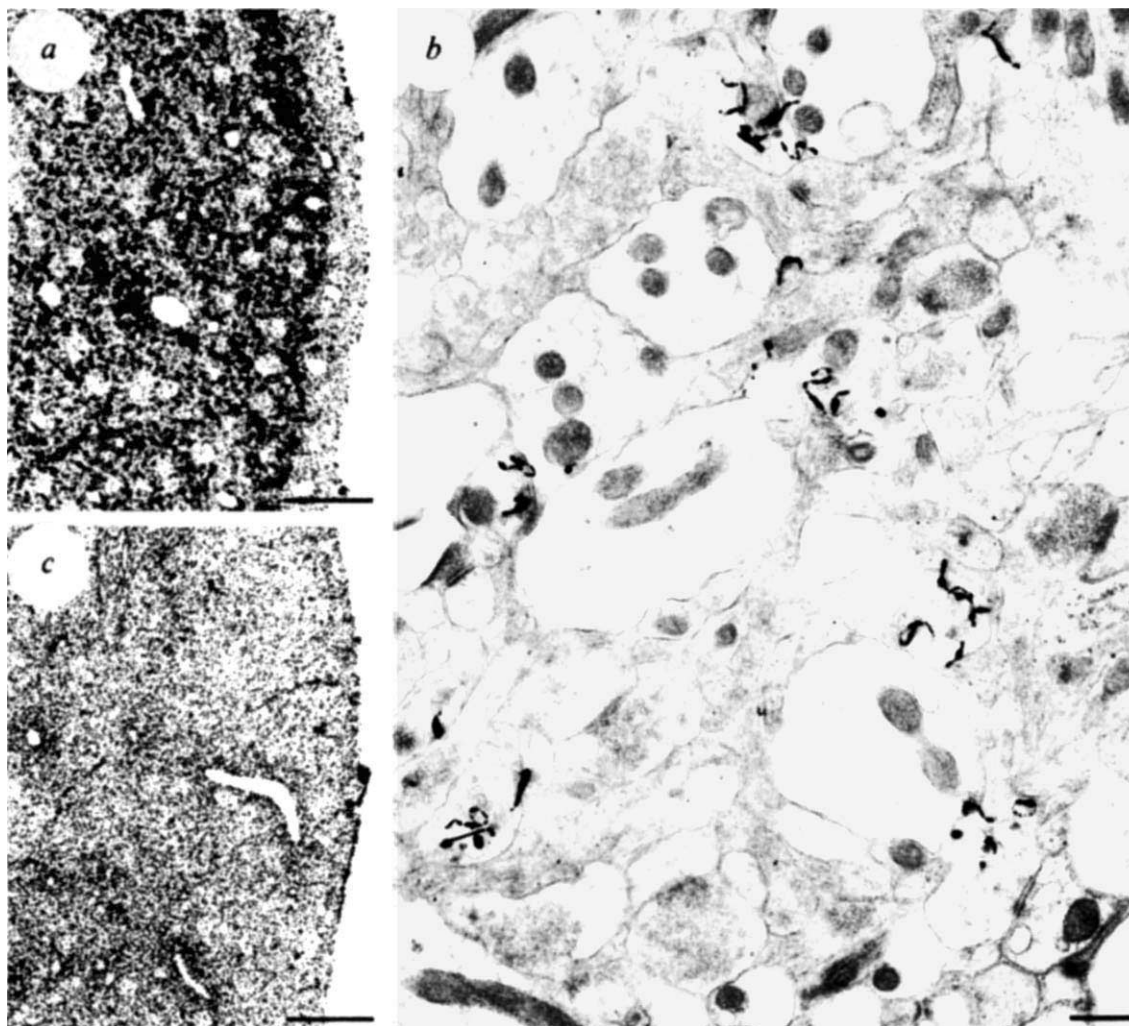


Fig. 1 Light (*a, c*) and electron microscope (*b*) autoradiographs of paraventricular neostriatum, after intraventricular instillation of ^3H -DA *in vivo* followed by rapid double perfusion-fixation with glutaraldehyde and osmium tetroxide solutions. *a*, ^3H -DA was administered together with non-radioactive serotonin. Small and dense aggregates of silver grains, disseminated throughout the neuropil, signal the existence of axonal varicosities having taken up and retained the tracer. *b*, All accumulations of silver grains correspond to axonal enlargements containing synaptic vesicles. Some of these varicosities are seen in synaptic junction with dendritic profiles. *c*, ^3H -DA was administered 15 d after destruction of nigrostriatal dopaminergic neurones by 6-hydroxydopamine. The silver grain aggregates are no longer visible, which demonstrates their exclusive dopaminergic axonal origin. *a*, And *c*, D-19 development after 15 and 30 d of exposure, respectively. Scale bars, 50 μm . *b*, Paraphenylenediamine development after 30 d of exposure. Scale bar 0.5 μm .

between results of *in vivo* and *in vitro* experiments could be due to altered properties of severed serotonergic axon terminals incubated *in vitro*, and/or insufficient tissue concentration of ^3H -DA *in vivo* to induce interspecific labelling of serotonergic nerve endings.

Definite proof of the specific identity of the ^3H -DA-accumulating nerve endings in neostriatum was provided by autoradiographic data obtained after 6-hydroxydopamine lesioning²⁰ of the nigrostriatal dopaminergic system. In these rats, the neuronal cell bodies giving rise to the dopaminergic innervation of neostriatum, located in groups A-9 and A-10 of zona compacta and area ventralis tegmenti²¹, were selectively destroyed by local microinjection of the drug (8 μg in 4 μl of saline), 15 d before ^3H -DA administration. In agreement with previous reports, this pretreatment resulted in the conspicuous disappearance of all histofluorescent perikarya in substantia nigra, accompanied by a 90% decrease of endogenous DA content and total loss of histofluorescence in neostriatum²²⁻²⁴. Its consequence in light microscope autoradiographs of neostriatum was the virtual absence of any ^3H -DA

accumulations, confirming their dopaminergic axonal origin (Fig. 1*c*).

It seems likely that the greater diffusion of ^3H -DA than other biogenic amines during standard histological processing reflects some basic difference in the binding of these molecules inside their respective axon terminals²⁵. In fact, by comparison with noradrenergic nerve endings elsewhere in brain, the dopaminergic terminals of neostriatum are known to exhibit less distinct endogenous histofluorescence when visualised with the Falck-Hillarp technique²⁶; they also show a linear fluorescence-concentration relationship, as if some endogenous DA and/or its fluorophore were extracted from 'granules' or even from axoplasm during histochemical processing²⁷. Furthermore, greater amounts of endogenous or exogenous DA than NA are recovered in supernatant when synaptosomes or particulate fractions from regions containing these amines are being prepared by differential centrifugation^{25,28-30}. Such findings have already led to the suggestion that a significant proportion of the biogenic amine within dopaminergic nerve terminals is located extravesicularly, in a weakly bound and readily releasable

pool³¹. This same hypothesis could also explain why the cytochemical demonstration of small dense-core vesicles in neostriatal axon terminals has only been possible after loading with exogenous NA or false transmitter³²⁻³⁵, but has always failed with DA even after tissue incubation in high concentrations of this amine³².

The mechanisms by which histoprocessing by vascular perfusion permits the retention of loosely bound DA in neostriatal dopaminergic nerve endings remain to be elucidated. Preliminary experiments suggest that comparable results may be obtained after successive perfusion of both glutaraldehyde and osmium fixatives and subsequent dehydration and embedding carried out by immersion of tissue slices in the usual manner. It will be of interest to determine whether factors such as the velocity of the double fixation achieved by vascular perfusion, and/or the preservation of unbroken cell membranes throughout glutaraldehyde and osmium fixations, contributes in maintaining ³H-DA *in situ*.

In any event, the present study offered the first opportunity to examine dopaminergic nerve endings of neostriatum in conditions allowing not only their unequivocal identification, but also adequate preservation of their ultrastructural features. Until now, examination of some 100 sectional profiles of dopaminergic axonal varicosities in the paraventricular zone of neostriatum has shown that these are relatively small, averaging 0.5 µm in diameter (0.3–0.9 µm). Most contain round and clear small synaptic vesicles (30–50 nm), usually accompanied by one or more mitochondria; only a few exhibit large dense-core vesicles (80–120 nm). Note also that a fair number of ³H-DA-labelled varicosities may be seen in close apposition to small nerve cell bodies or in symmetrical synaptic junction with dendritic processes (Fig. 1b). Further elucidation of the fine structural features of neostriatal dopaminergic axonal varicosities should help to resolve present controversies regarding the junctional versus non-junctional relationships of these nerve terminals³⁵.

In regions of the brain known to receive a mixed dopaminergic and noradrenergic innervation (for example hypothalamic paraventricular nucleus and median eminence), the distribution of axon terminals detected in light microscope autoradiographs suggests that both types of catecholaminergic nerve endings are simultaneously labelled with ³H-DA after double perfusion of glutaraldehyde and osmium fixatives. Selective cytolytic removal of noradrenergic afferents and/or pharmacological inhibition of their uptake systems, might therefore prove necessary for achieving a specific autoradiographic identification of dopaminergic nerve endings within such parts of the nervous system.

We thank C. de Montigny and T. A. Reader for their histo- and biochemical analyses, K. C. Watkins, S. Garcia and D. Cyr for technical assistance, and J. Manson for reviewing the manuscript. These studies were supported by MRC grant MT-3544 (to L.D.), funds from CAFIR and the Magali Ducharme and Joseph Rhéaume foundations at the Université de Montréal, and a CRSQ fellowship (to O.B.).

Received 7 December 1979; accepted 12 February 1980.

- Wolfe, D. E., Potter, L. T., Richardson, K. C. & Axelrod, J. *Science* **138**, 440–442 (1962).
- Calas, A. & Segu, L. *J. Microscopic Biol. Cell* **27**, 249–252 (1976).
- Descarries, L. & Droz, B. *C. r. heb. Séanc. Acad. Sci., Paris* **266**, 2480–2482 (1968).
- Parizek, J., Hassler, R. & Bak, I. *J. Z. Zellforsch.* **115**, 137–148 (1971).
- Sotelo, C. *J. ultrastruct. Res.* **36**, 824–841 (1971).
- Scott, D. E., Sladek, J. R., Knigge, K. M., Krobisch-Dudley, G., Kent, D. L. & Sladek, C. D. *Cell Tissue Res.* **166**, 461–473 (1976).
- Halász, N., Ljungdahl, Å. & Hökfelt, T. *Brain Res.* **154**, 253–271 (1978).
- Aghajanian, G. K. & Bloom, F. E. *J. Pharmac. exp. Ther.* **156**, 407–416 (1967).
- Calas, A. *et al. Brain Res.* **118**, 1–13 (1976).
- Arлуison, M., Agid, Y. & Javoy, F. *Neuroscience* **3**, 675–683 (1978).
- Des Rosiers, M. H. & Descarries, L. *C. r. heb. Séanc. Acad. Sci., Paris* **287**, 153–156 (1978).
- Descarries, L., Bosler, O. & Des Rosiers, M. H. *Neurosci. Lett. Suppl.* **3**, S 232 (1979).
- Larra, F. & Droz, B. *J. Microscopic Biol.* **9**, 845–880 (1970).
- Descarries, L. & Droz, B. *J. Cell Biol.* **49**, 385–399 (1970).
- Swanson, L. W. & Hartman, B. K. *J. comp. Neurol.* **163**, 467–506 (1975).
- Glowinski, J. & Iversen, L. L. *J. Neurochem.* **13**, 655–669 (1966).
- Hamberger, B. *Acta physiol. scand. Suppl.* **295**, 1–56 (1967).
- Fuxe, K. & Ungerstedt, U. *Histochemia* **13**, 16–28 (1968).
- Berger, B. & Glowinski, J. *Brain Res.* **147**, 29–45 (1978).
- Ungerstedt, U. *Eur. J. Pharmac.* **5**, 107–110 (1968).

- Andén, N.-E., Carlsson, A., Dahlström, A., Fuxe, K. & Hillarp, N.-Å. *Life Sci.* **3**, 523–530 (1964).
- Hökfelt, T. & Ungerstedt, U. *Brain Res.* **60**, 269–297 (1973).
- Sotelo, C., Javoy, F., Agid, Y. & Glowinski, J. *Brain Res.* **58**, 269–290 (1973).
- Agid, Y., Javoy, F., Glowinski, J., Bouvet, D. & Sotelo, C. *Brain Res.* **58**, 291–301 (1973).
- Gutman, Y. & Weil-Malherbe, H. *J. Neurochem.* **14**, 619–625 (1967).
- Fuxe, K. *Acta physiol. scand.* **64**, Suppl. 247, 37–85 (1965).
- Einarsson, P., Hallman, H. & Jonsson, G. *Med. Biol.* **53**, 15–24 (1975).
- Glowinski, J. & Iversen, L. L. *Biochem. Pharmac.* **15**, 977–987 (1966).
- Philippu, A. & Heyd, W. *Life Sci.* **9**, 361–373 (1970).
- Nose, T., Segawa, T. & Takagi, H. *Japan J. Pharmac.* **22**, 867–869 (1972).
- Glowinski, J., Besson, M. J., Chéramy, A., Javoy, F. & Thierry, A. M. *Rev. Neurol., Paris* **127**, 1–22 (1971).
- Hökfelt, T. & Jonsson, G. *Histochemia* **16**, 45–67 (1968).
- Hökfelt, T. *Z. Zellforsch.* **91**, 1–74 (1968).
- Tennyson, V. M., Heikkilä, R., Mytilineou, C., Côté, L. & Cohen, G. *Brain Res.* **82**, 341–348 (1974).
- Arлуison, M., Agid, Y. & Javoy, F. *Neuroscience* **3**, 657–673 (1978).

Role of NK cells in tumour growth and metastasis in *beige* mice

James E. Talmadge*†, Kenneth M. Meyers‡, David J. Prieur* & Jean R. Starkey*

* Department of Veterinary Microbiology and Pathology, Washington State University

‡ Department of Veterinary and Comparative Anatomy, Pharmacology and Physiology, Washington State University, Pullman, Washington 99164

Although natural killer (NK) cells are thought to give the host a spontaneous resistance against tumours and have been postulated to act *in vivo* as surveillance cells¹, definitive data in support of these hypotheses has not been obtained. Recently the *beige* (*bg*) mouse, a morphological homologue of the human Chediak-Higashi (CH) syndrome²⁻⁴, was shown to be deficient in NK activity⁵. Specifically, spleen cells of *bg* mice were demonstrated to be incapable of *in vitro* natural cytotoxicity against tumour cells⁵. We report here that a tumour line, modified to be sensitive to NK cytotoxicity by *in vitro* culture, demonstrated *in vivo* an increased growth rate, faster induction time and an increased metastatic capability in *bg* compared to control mice. This was not found with a tumour line insensitive to NK activity (without *in vitro* culture). *In vivo* activation of NK cells in *bg* and control mice resulted in a decrease in tumour growth rate and metastatic frequency. These results demonstrate that NK cells have an important function in the host's control of tumour growth and metastasis.

NK cells seem to be pre-T cells that do not require thymic maturation. This could account for their presence in athymic nude mice⁶. The level of NK cell activity is genetically controlled⁷, age-dependent⁸ and may be augmented by numerous agents, including infection with lymphocytic choriomeningitis virus (LCMV)⁹. Induced NK-cell activity peaks 3 d after injection of LCMV and decreases to normal levels by day 8 (ref. 10). Interferon levels directly parallel the rise and decrease of NK activity¹¹ suggesting that the increased NK-cell activity in LCMV infected mice is mediated by the viral induction of interferon¹⁰. The sensitivity of tumour cell lines to NK-cell activity varies directly with the length of time they have been cultured *in vitro*, and *in vivo* passaged cell lines are often insensitive to NK cytotoxicity⁸. At least 3 weeks of *in vitro* culture are required for a tumour line to develop an increased sensitivity to NK cytotoxicity¹².

Growth parameters and metastatic abilities of the malignant melanoma tumour B16 were studied in syngeneic normal (+/*bg*) and *bg* (*bg/bg*) C57BL/6 mice. B16 tumours which were passaged *in vivo* and cultured *in vitro* only a short time were insensitive to NK lysis (Table 1A). However, increased *in vitro* culture time resulted in the acquisition of a sensitivity to NK lysis

† To whom correspondence should be addressed at: Frederick Cancer Research Center, Biology of Metastasis Section, PO Box B, Frederick, Maryland 21701.

by the tumour cells (Table 1A). Table 1 also demonstrates that spleen cells from normal mice exhibit NK activity but that spleen cells from *bg* mice do not exhibit unstimulated NK activity, confirming a previous report⁵. LCMV induction of NK activity markedly enhanced cytotoxicity against B16 cells cultured *in vitro* for 32 d as well as against B16 cells cultured *in vitro* for only one day that were refractory to unstimulated NK activity (Table 1A).

The injection of LCMV into 6–8-week-old mice increased NK activity in both normal and *bg* mice. In normal mice NK activity peaked on day 3 and returned to normal by day 6 following LCMV injection. The *bg* mice expressed NK activity 24 h after LCMV injection. This activity peaked on day 5, and decreased more slowly than in normal mice, a finding in agreement with that of Kiessling¹².

Tumours insensitive to NK lysis required the same time to achieve 100% tumour induction and grew to essentially the same mean tumour volumes in normal and *bg* mice (Table 2). However, tumour cells sensitive to NK lysis had longer tumour induction times, slower growth rates (data not shown) and produced smaller tumours in normal compared to *bg* mice (Table 2). The parameters of tumour growth for both the NK-insensitive and -sensitive tumour lines were reduced when injected into LCMV-infected normal or *bg* mice (Table 2). Spontaneous metastasis was assessed following subcutaneous tumour growth in a pinna, followed by resection of the ear on day 21, and necropsy of mice at the onset of morbidity¹³. Approximately the same number of metastases were found in normal and *bg* mice injected with NK-insensitive tumour cells

Table 1 Sensitivity of B16 cultures to NK lysis and activation of NK cells in *beige* (*bg/bg*) and normal (+/*bg*) mice

A Effect of various intervals of *in vitro* culture on the sensitivity of B16 cells to NK cell lysis from the spleens of *bg* and normal mice

Length of B16 culture <i>in vitro</i> (d)	Per cent cytotoxicity						LCMV-activated effector cells		
	Effector cells <i>bg/bg</i> (Effector to target ratio)			Effector cells +/ <i>bg</i> (Effector to target ratio)			+/ <i>bg</i> (Effector to target ratio)		
	200:1	100:1	50:1	200:1	100:1	50:1	200:1	100:1	50:1
1	0.0	0.0	0.0	5.6	0.0	0.0	52	52	38
1	0.0	0.0	0.0	0.0	0.0	0.0	40	44	32
1	0.0	0.0	0.0	4.6	0.0	0.0	50	44	28
32	2.0	0.0	0.0	14.0	8.8	2.0	64	64	42
32	3.3	0.0	0.0	14.0	8.6	5.0	68	56	34
32	2.1	0.0	0.0	20.0	10.0	6.6	66	56	42

B Time of appearance of activated NK cells in the spleens of *bg* and normal mice following LCMV injection

Days postinfection	Per cent cytotoxicity			
	Effector cells +/ <i>bg</i>		Effector cells <i>bg/bg</i>	
	B ₁	B ₂	B ₁	B ₂
0 Control	8	10	4	3
1	28	34	6	12
2	28	34	16	20
3	40	42	28	36
4	12	14	32	42
5	14	16	38	46
6	8	8	26	28

A, Cultured target cells were labelled with ⁵¹Cr (2 × 10⁷ cells were incubated with 200 μCi ⁵¹Cr for 40 min at 37 °C), washed, allowed to attach to the wells of a microtest II dish (5,000 cells per well) for 4 h, and washed again. Complete medium (200 μl) containing effector cells from the spleens of: (1) *bg*; (2) normal; or (3) normal mice infected 3 d previously with 1–2 × 10⁴ plaque-forming units (PFU) of LCMV were then added. Following 6 h incubation at 37 °C 100-μl aliquots of the supernatant were removed and counted in a γ counter. Spontaneous release was obtained from wells incubated without effector cells and was always less than 10% of total release obtained with 2% SDS (450 cpm/6,000 cpm total release). NK activity is expressed as the per cent cytotoxicity = (test c.p.m. – spontaneous c.p.m.)/(total c.p.m. – spontaneous c.p.m.).

B, NK lysis activity was assayed as in experiment A, using B16 target cells which had been cultured *in vitro* for 32 d. Mice that were 6–8 weeks of age were injected with 1–2 × 10⁴ PFU of LCMV on consecutive days and assayed on the same day using an effector: target ratio of 100:1.

Table 2 Role of NK cells in modifying subcutaneous growth and metastasis of B16 tumour cells

Time required to reach 100% tumour induction, or the percentage of animals with tumours 21 days following tumour injection

Mouse genotype	LCMV	B16 cells cultured for:	
		5 days	32 days
<i>bg/bg</i>	–	13 days	14 days
+/ <i>bg</i>	–	14 days	45%
<i>bg/bg</i>	+	16 days	17 days
+/ <i>bg</i>	+	80%	80%
Mean tumour size ± s.e.m. 21 days following tumour cell injection (mm ³)			
Mouse genotype	LCMV	B16 cells cultured for:	
		5 days	32 days
<i>bg/bg</i>	–	105 ± 26	93 ± 22
+/ <i>bg</i>	–	175 ± 57	4.1 ± 1.2
<i>bg/bg</i>	+	81 ± 36	42 ± 19
+/ <i>bg</i>	+	25 ± 15	5.4 ± 1.4
Incidence of pulmonary metastases (mean number of metastases ± s.e.m. per mouse)			
Mouse genotype	LCMV	B16 cells cultured for:	
		5 days	32 days
<i>bg/bg</i>	–	5.4 ± 0.95	7.1 ± 0.88
+/ <i>bg</i>	–	6.1 ± 0.63	0 ± 0
<i>bg/bg</i>	+	0 ± 0*	0 ± 0*
+/ <i>bg</i>	+	0 ± 0*	0 ± 0

The mice received 10⁵ cells in a 0.05-ml aliquot subcutaneously in one ear. Mice with LCMV had received 1–2 × 10⁴ PFU 2 days previously. The tumour size was measured in three dimensions after 21 days of growth and the volume calculated¹⁷. The tumour bearing ear was amputated at this time. 15 days later the mice were necropsied and the number of spontaneous metastases counted. *n* = 10 mice.

* Several of these mice had extensive peripheral lymph node tumour growth.

Table 3 Role of host NK activity and tumour sensitivity to NK lysis in modifying the incidence of metastases following intravenous injection of B16 tumour cells

A	Mouse genotype	LCMV	Mean no. lung metastases ± s.e.m. per mouse	
			B16 tumour cells cultured for:	
			5 days	32 days
	<i>bg/bg</i>	–	6.7 ± 0.88	8.4 ± 2.71
	+/ <i>bg</i>	–	9.2 ± 1.48	0 ± 0
	<i>bg/bg</i>	+	3.1 ± 0.5	4.2 ± 0.73*
	+/ <i>bg</i>	+	3.3 ± 0.31	0 ± 0
B			2 days	40 days
	<i>bg/bg</i>	–	17.4 ± 2.27	16.8 ± 2.78
	+/ <i>bg</i>	–	11.0 ± 1.45	0 ± 0

Mice received 0.10 ml of a single cell suspension of B16 cells (experiment A, 20,000 cells; experiment B, 50,000 cells). These cells had been cultured for 5 or 32 days (experiment A), and 2 or 40 days (experiment B). The mice injected with LCMV received 1–2 × 10⁴ PFU of LCMV 2 days before tumour cell injection. After 15 days the mice were necropsied and the number of pulmonary metastases counted using a dissecting microscope. *n* = 10 mice.

* These mice had five extra pulmonary metastases, three ovarian and two renal.

(Table 2). However, the incidence of spontaneous metastasis was markedly reduced with NK lysis sensitive tumour cells injected into normal mice compared to *bg* mice. The metastatic incidence was greatly reduced in mice with LCMV-activated NK cells bearing NK-insensitive tumours in normal or *bg* mice and NK-sensitive tumours in *bg* mice.

The metastatic abilities of NK-sensitive and -insensitive B16 cells were also studied using the lung colony assay (Table 3). No difference was found in the experimental metastatic ability of NK-insensitive tumour cells in normal or *bg* mice (Table 3). However, NK-sensitive tumour cells were metastatic using the lung colony assay in *bg* mice but not in normal mice (Table 3). Metastatic ability seemed to correlate with the tumour cells' sensitivity to NK lysis and with the level of NK activity. This conclusion was further substantiated using mice of both genotypes infected with LCMV. These mice had fewer lung colony foci following injection with either NK lysis sensitive or insensitive tumour cells compared to mice without activated NK cells, except for NK-sensitive cells injected into normal mice where they were already non-metastatic.

A role for NK cells in metastasis was initially suggested by the observation that transplanted tumours rarely metastasise in nude mice¹⁴ which have high levels of NK activity¹⁵. Also there is a direct correlation between the level of NK activity in different mouse strains and the short-term survival of tumour cells labelled with ¹²⁵I-iododeoxyuridine using a lung colony assay¹⁶. An NK-sensitive tumour injected into *bg* mice (deficient in normal NK activity) is induced faster, grows more rapidly and has more metastases than the same tumour in normally NK-constituted mice. These differences were not present with a tumour insensitive to NK cells. Normal and *bg* mice with

LCMV-activated NK cells, however, exhibited a reduced growth rate and reduced metastatic incidence of both NK lysis sensitive and insensitive tumour cells. These results suggest an important role for NK cells in the host's control of malignant disease.

We thank Dr R. Welsh and Dr R. Kiessling for helpful discussions and Dr Welsh also for providing LCMV. This study was supported by NIH grant RR 00515 and by American Cancer Society Institutional support grant no. IN-119. D.J.P. is the recipient of an NIH Research Career Development Award (RR00003).

Received 13 November 1979; accepted 7 January 1980.

- Haller, O., Hansson, M., Kiessling, R. & Wigzell, H. *Nature* **270**, 609 (1977).
- Lutzner, M. A., Lowrie, C. T. & Jordan, H. W. *J. Hered.* **58**, 299–300 (1967).
- Windhorst, D. B. & Padgett, G. J. *Invest. Derm.* **60**, 529–537 (1973).
- Klebanoff, S. J. & Clark, R. A. in *The Neutrophil: Function and Clinical Disorders* (North-Holland, Amsterdam, 1978).
- Roder, J. & Duwe, A. *Nature* **278**, 451–453 (1979).
- Herberman, R. B., Nynn, M. E. & Lavrin, D. H. *Int. J. Cancer* **16**, 230–239 (1975).
- Petranyi, G. G., Kiessling, R., Povey, S., Klein, G., Herzenberg, E. & Wigzell, H. *Immunogenetics* **3**, 15–28 (1976).
- Kiessling, R., Klein, E. & Wigzell, H. *Eur. J. Immun.* **5**, 112–117 (1975).
- Welsh, R. M. & Zinkernagel, R. M. *Nature* **268**, 646–648 (1977).

- Gidlund, M., Örn, A., Wigzell, H., Senik, A. & Gresser, I. *Nature* **273**, 759–761 (1978).
- Welsh, R. M. *J. exp. Med.* **163**, 181 (1978).
- Becker, S., Kiessling, R., Lee, N. & Klein, G. *J. natn. Cancer Inst.* **61**, 1495–1498 (1978).
- Fidler, I. J., Gersten, D. M. & Kripke, M. L. *Cancer Res.* **39**, 3816–3821 (1979).
- Gershwin, M. E., Ikeda, R. M., Kawakami, T. G. & Owens, R. B. *J. natn. Cancer Inst.* **58**, 1455–1460 (1977).
- Herberman, R. B. in *The Nude Mouse in Experimental and Clinical Research* (Academic, New York, 1978).
- Riccardi, C., Puccetti, P., Santoni, A. & Herberman, R. B. *J. natn. Cancer Inst.* **63**, 1041–1045 (1979).
- Dethlefsen, L. A., Prewitt, J. M. S. & Mendelsohn, M. L. *J. natn. Cancer Inst.* **40**, 389–405 (1968).

Low natural *in vivo* resistance to syngeneic leukaemias in natural killer-deficient mice

Klas Kärre, Gunnar O. Klein, Rolf Kiessling, George Klein & John C. Roder*

Department of Tumor Biology, Karolinska Institutet, S-104 01 Stockholm 60, Sweden

Natural killer (NK) cells, which occur at high levels in normal non-immune individuals of several species, have the capacity to lyse certain tumour cells *in vitro*, and have been proposed as a first level of defence against tumour growth *in vivo*^{1–3}. Evidence for this comes mainly from the murine system, where certain F₁ hybrid or T-cell deficient mice with high NK activity resist growth of transplanted tumours better than do normal syngeneic controls with lower NK activity^{4–6}. Recent studies of the beige mutation on the C57BL mouse strain background, previously proposed as an animal homologue of the human Chediak-Higashi syndrome^{7,8}, have revealed a new model for analysing anti-tumour effector mechanisms in the intact syngeneic host^{9,10}. The recessive beige gene (*bg*¹), which affects melanosome and granulocyte lysosome functions^{11,12}, also causes a profound (but not total) depression of NK activity whereas other anti-tumour effector mechanism mediated by T cells and macrophages remain relatively intact¹³. It was important, therefore, to test the tumour susceptibility of beige mice as a first direct test of the hypothesis that NK cells are involved in surveillance against neoplasia. In partial confirmation of this hypothesis, we report here that failure of small threshold doses of one virally and one chemically induced transplantable leukaemia to grow out in syngeneic mice partly depends on a rapidly acting host defence mechanism. This mechanism is deficient in *bg/bg* mice, which develop palpable, progressively growing tumours faster and at a higher frequency than do phenotypically normal *+/bg* littermates.

The benzyrene-induced EL-4 and the Rad-LV-induced P-52-127-166 leukaemias of C57BL/6 origin, are both highly resisted by certain C57BL F₁ hybrids that have high NK activity *in vitro*^{14,15}. Also, low-dose inocula of the *in vivo* maintained ascites lines of the two leukaemias behaved similarly; subcutaneous (s.c.) inoculation in a transplantation test yielded higher tumour take incidence in *bg/bg* mice than in heterozygous littermates (Fig. 1). The differences were most striking in

the early phase of the observation period, with a majority of the total tumour takes in *bg/bg* registered within 2 weeks (68% with EL-4, 69% with P-52-127-166) whereas only a small part of the tumours in *+/bg* mice were palpable at this time (31% with EL-4, 14% with P-52-127-166). Thus, in addition to the increased incidence of takes in mutant mice, the progressively growing tumours appeared with a shorter latency than in controls. This was also reflected in the larger mean tumour diameter by 2 and 3 weeks and the earlier deaths among tumour-bearing *bg/bg* than *+/bg* mice (Fig. 1 legend).

The outcome of a long-term transplantation test can be influenced by a variety of host factors. However, most of the tumours in *bg/bg* mice had appeared after the short period of 2 weeks, after which the tumour incidence curves of *bg/bg* and *+/bg* became virtually parallel. This indicated that the defect in natural resistance of beige mice involved early events after tumour inoculation. Further support for this was obtained in a test system where the short-term survival of intravenously (i.v.) injected ¹²⁵I-iododeoxyuridine (IUdR) labelled leukaemia cells is monitored, previously proposed as a most appropriate *in vivo* assay for natural rejection mechanisms^{16,17}, including NK cells^{18,19}. The heterozygous control mice eliminated P-52-127-166 ascites cells more efficiently than did beige mice, measured either as total, pulmonary or splenic radioactivity retained 18 h after i.v. injection (Fig. 2), although in the spleen, differences were already detectable after 4 h (Fig. 2 legend). The results suggest that *bg/bg* mice may lack an important mechanism for rapidly eliminating tumour cells, for example, as blood-borne metastases in the lungs and spleen.

Our findings in *in vivo* tests correlate well with the *in vitro* measured splenic NK activity against both tumours tested. EL-4 or P-52-127-166 cells had a low, but significant, sensitivity to *+/bg* spleen cell cytotoxicity (5–10% specific lysis) whereas *bg/bg* spleen consistently gave values below 3% specific lysis (Fig. 3).

Augmentation of the low NK levels against EL-4 and P-52-127-166 occurred in mice receiving the low tumour cell dose used in the transplantation tests (Fig. 3). The NK activities of *+/bg* as well as *bg/bg* mice were increased, although relative differences between the groups were maintained. In both tumour-inoculated and control groups, *bg/bg* mice gave values corresponding to a ninefold dilution of *+/bg* splenocytes in terms of lytic units. As has been shown with larger doses of tumour cells²⁰, the augmented cytotoxicity observed here was mediated by non-adherent cells (Fig. 3) which were relatively resistant to treatment with anti-Thy 1.2 plus complement, and declined in activity after day 3 to values comparable to those in non-inoculated control mice (unpublished observation). Such augmented NK activity, induced during the transplantation

* Present address: Department of Microbiology and Immunology, Queen's University, Kingston, Canada.

tests, may account for *in vivo* resistance against tumour cells which show very low susceptibility to *in vitro* lysis by NK cells from normal mice.

Several different phenotypic manifestations of the beige mutation may be partly responsible for the low natural resistance against transplanted tumours in *bg/bg* mice observed here. However, our results from the two *in vivo* test systems suggest that early events are of major importance for the final outcome of tumour growth; this would argue against adaptive mechanisms requiring proliferative expansion of specific clones, as in a cytotoxic T-cell response. Furthermore, the early more efficient *in vivo* elimination of leukaemia cells in *+/bg* mice was paralleled by a greater cell-mediated cytotoxicity *in vitro* against the same tumours which was independent of adherent cells. This makes the involvement of monocytes or macrophages, proposed as responsible for natural surveillance in other systems^{21,22}, less likely. Finally, *bg/bg* mice generate highly efficient T-killer cells *in vivo* after injection of P815 tumour cells, and anti-tumour effects mediated by macrophages are also normal¹³. Although we cannot exclude the possibility that mechanisms which are

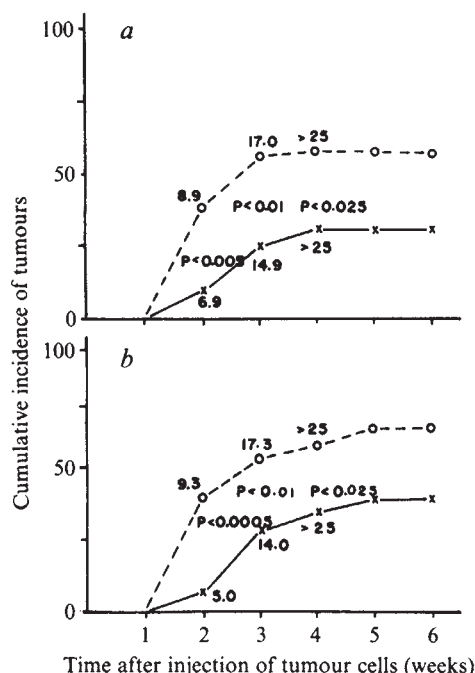


Fig. 1 Cumulative tumour take incidences in *+bg* and *bg/bg* mice. EL-4 or P-52-127-166 cells were collected from the ascitic fluid of C57BL/6 mice (in which the tumours were propagated *in vivo* in our animal colony) and washed once in balanced salt solution (BSS). EL-4 (5×10^3) or P-52-127-166 (10^3) cells were then injected s.c. in the back of *+bg* or *bg/bg* littermates, obtained by crossing C57BL/6 *+bg* and C57BL/6 *bg/bg* breeders (originally purchased from Jax Laboratories and maintained by continuous brother-sister mating according to a backcross-intercross scheme in our animal colony). The challenging cell doses were chosen from pretitration experiments so that tumours would grow progressively only in a fraction (30–40%) of syngeneic *+bg* mice. By this procedure, the inoculum would be large enough to be able to establish a growing tumour mass, but still so small that it would allow detection of any impairment of natural resistance in *bg/bg* mice by an increased incidence of tumours. Curves and P values (χ^2 -test) compare the tumour take incidence in *+bg* (x) and *bg/bg* (o) at weekly intervals, with figures at each time point referring to mean tumour diameter (out of three perpendicular measurements) in mm. a), EL-4 (four experiments with a total of 52 *+bg* and 38 *bg/bg* mice tested): out of all *+bg* mice that developed progressively growing tumours, 6% had died by week 3 and 65% by week 4; the corresponding figures for *bg/bg* mice were 22% by week 3 and 95% by week 4. b), P-52-127-166 (6 experiments with a total of 70 *+bg* and 56 *bg/bg* mice tested): out of all *+bg* mice that developed progressively growing tumours, 0 had died by week 3, 21% by week 4 and 69% by week 5; the corresponding figures for *bg/bg* mice were 6%, 57% and 89%, respectively. No new palpable tumours appeared after 5 weeks (mice were kept for 8–12 weeks after tumour injection). Preliminary experiments with 10-fold higher tumour cell doses have yielded 80–90% takes in beige as well as control animals, although latency and survival data show faster tumour growth in *bg/bg* mice.

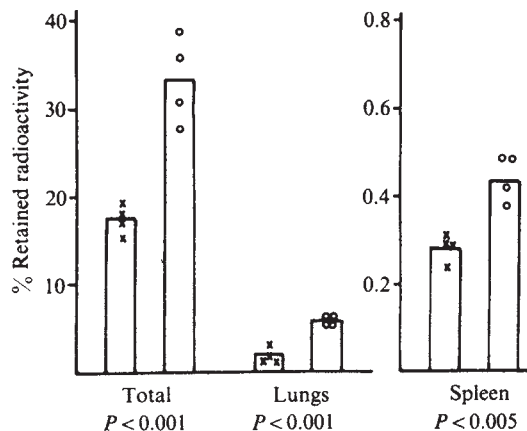


Fig. 2 *In vivo* elimination of $^{125}\text{IUdR}$ -labelled leukaemia cells in *+bg* and *bg/bg* mice. P-52-127-166 ascites cells were labelled by intraperitoneal injection of $50 \mu\text{Ci}$ ^{125}I -labelled 5-iodo-2'-deoxyuridine ($>5 \text{ Ci per mg}$, Radiochemical Centre) 3 days after passage of 10^7 cells. After 6 h, cells were collected by peritoneal washes, counted, washed four times in BSS and injected in a lateral tail vein of *+bg* and *bg/bg* mice (1×10^6 cells per mouse). Groups of four or five animals were killed after 4 and 18 h, and the spleens and lungs were removed. Whole body and organ counting were carried out in an Intertechnique γ spectrometer. Symbols represent total (whole body + spleen + lungs), splenic and pulmonary retained radioactivity (calculated as percentage of total injected) in individual *+bg* (x) and *bg/bg* (o) mice after 18 h. Vertical columns represent arithmetic means, and P values refer to statistical analysis by a Student's *t*-test. Four hours after injection, the corresponding values were as follows (mean and range): *bg/bg*, 83(71–89); *+bg*, 82 (77–80); lungs: *bg/bg*, 61 (57–69); *+bg*, 54 (49–60); $P < 0.1$; spleen: *bg/bg*, 0.52 (0.47–0.56); *+bg*, 0.41 (0.33–0.52), $P < 0.05$.

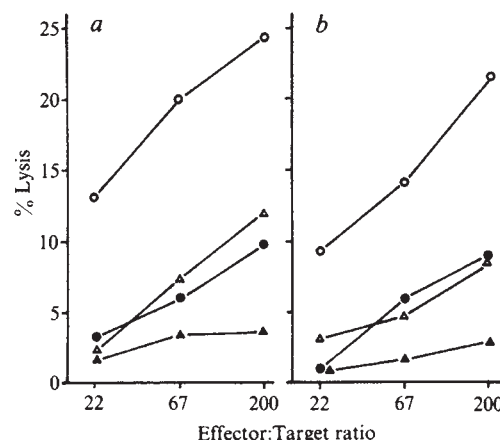


Fig. 3 Splenic NK activity of normal or tumour-inoculated *+bg* and *bg/bg* mice. Three days after s.c. inoculation of BSS or 10^3 P-52-127-166 cells, mice were killed and spleen cell suspensions prepared in complete RPMI with 10% fetal calf serum. Cells from three pooled spleens were incubated in a nylon wool column²⁶ and thereafter used as effectors in a standard 8-h ^{51}Cr release assay (as previously described) against EL-4 (a) or P-52-127-166 (b) cells. Values represent the mean of triplicate samples (triplicate variation $<5\%$ of mean value). Spontaneous release was 11.2% for EL-4 and 15.1% for P-52-127-166. Unfractionated control cells were in all cases slightly less efficient than nylon wool-passed cells as cytotoxic effectors (not shown). ●, *+bg* (BSS); ○, *+bg* (10^3 P52-127-166); ▲, *bg/bg* (BSS); △, *bg/bg* (10^3 P52-127-166).

unaffected by the beige mutation contribute to natural resistance, our results indicate that NK cells have a major role in the cell-mediated *in vivo* defence against transplanted leukaemias observed in syngeneic phenotypically normal (*+bg*) mice. This has important implications for the evaluation of the influence of NK cells on primary tumour formation and growth in the normal autochthonous host. If the high NK-cell activity of nude mice is the explanation for their low or 'normal' incidence of tumours and their ability to resist growth of certain murine as well as human transplanted tumours, it is clear that a double mutant combining the beige and the nude defects would be a useful tool

for future research. Such a defect is not lethal, and we have recently been able to produce double mutant mice by a series of crosses and intercrosses between C57BL/6 *bg/bg* and C57BL/6+/*nu* mice. These C57BL/6 *bg/bg*, *nu/nu* mice are now being characterised immunologically.

In conclusion, the experimental model provided by the beige mutation provides strong additional support for the notion that NK cells may be involved in anti-tumour surveillance *in vivo*. Conclusive evidence should come from further studies of primary induced or spontaneous tumours in beige mice. In this context, it is interesting that humans bearing the Chediak-Higashi gene have an even greater impairment (500-fold) in NK

function²³ and a corresponding increase in the incidence of a spontaneous lymphoproliferative disorder which is thought to be malignant²⁴.

During these studies, we became aware of a similar study in progress by Talmadge *et al.*²⁵ (*Nature*, this issue), showing faster growth and increased metastasis of *in vitro* cultured syngeneic melanoma cells in *bg/bg* mice.

We thank Dr Mona Hansson for helpful discussions, and Ms Maj-Lis Solberg and Ms Margareta Hagelin for technical assistance. This investigation was supported in part by grants 1R01 CA 26782-01 and 1R01 CA 14054-07 from the NCI, DHEW, and by grants from the Swedish Cancer Society.

Received 29 November 1979; accepted 27 February 1980.

- Kiessling, R. & Wigzell, H. *Immun. Rev.* **44**, 165–208 (1979).
- Herberman, R. B. & Holden, H. T. *Adv. Cancer Res.* **27**, 305–377 (1977).
- Haller, O., Hansson, M., Kiessling, R. & Wigzell, H. *Nature* **270**, 609–611 (1977).
- Petranyi, G. *et al. Immunogenetics* **3**, 15–28 (1976).
- Warner, N. L., Woodruff, M. F. & Burton, R. C. *Int. J. Cancer* **20**, 146–153 (1977).
- Kiessling, R., Petranyi, G., Klein, G. & Wigzell, H. *Int. J. Cancer* **17**, 275–281 (1976).
- Windhurst, D. B. & Padgett, G. J. *Invest. Derm.* **60**, 529–537 (1973).
- Blume, R. S. & Wolff, S. M. *Medicine, Baltimore* **51**, 247–280 (1972).
- Roder, J. C. & Duwe, A. K. *Nature* **278**, 451–453 (1977).
- Roder, J. C. *J. Immun.* **123**, 2168–2173 (1979).
- Gallin, J. I., Bujak, J. S., Palten, E. & Wolff, S. M. *Blood* **43**, 201–206 (1974).
- Elin, R. J., Edelin, J. B. & Wolff, S. M. *Infect. Immun.* **10**, 88–91 (1974).
- Roder, J. C., Lohmann-Matthes, M. -L., Domzig, W. & Wigzell, H. *J. Immun.* **123**, 2174–2181 (1979).
- Harmon, R. C., Clark, E. A., O'Toole, C. & Wicker, L. S. *Immunogenetics* **4**, 601–607 (1977).
- Klein, G., Klein, G. O., Kärre, K. & Kiessling, R. *Immunogenetics* **7**, 391–404 (1978).
- Carlson, G. A. & Wegman, T. G. *J. Immun.* **118**, 2130–2137 (1977).
- Carlson, G. A., Melnychuk, D. & Meeker, M. J. *Int. J. Cancer* **25**, 111–122 (1980).
- Riccardi, C., Puetti, P., Santoni, A. & Herberman, R. B. *J. natn. Cancer Inst.* **63**, 1041–1045, 1979.
- Herberman, R. B. *et al. Immun. Rev.* **44**, 43–70 (1979).
- Herberman, R. B., Nunn, M. E., Holden, H. T., Staal, S. & Djeu, J. Y. *Int. J. Cancer* **19**, 555–562 (1977).
- Greenberg, A. H. & Greene, M. *Nature* **264**, 356–369 (1976).
- Chow, D. A., Greene, M. I. & Greenberg, A. H. *Int. J. Cancer* **23**, 788–797 (1979).
- Roder, J. C. *et al. Nature* **284**, 553–555 (1980).
- Dent, P. B., Fish, L. A., White, J. F. & Good, R. A. *Lab. Invest.* **15**, 1634–1642 (1966).
- Talmadge, J. E., Meyers, K. M., Prieur, D. J. & Starkey, J. R. *Nature* **284**, 622–624 (1980).
- Julius, M. H., Simpson, E. & Herzenberg, L. A. *Eur. J. Immun.* **3**, 645–653 (1973).

Retinoids specifically enhance the number of epidermal growth factor receptors

Anton M. Jetten

Differentiation Control Section, Laboratory of Experimental Pathology, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20205

Retinoids elicit many biological and biochemical responses from cells *in vitro*^{1–5}. One widely used criterion for the responsiveness of cells to retinoids is inhibition of growth; retinoids reduce the saturation density and/or growth rate of many normal and tumorigenic cell lines^{2,6,7}. Propagation of eukaryotic cells has been demonstrated to be dependent on the presence of macromolecular growth factors such as epidermal growth factor (EGF)^{8–10}, which can stimulate proliferation of epithelial and fibroblastic cell lines. We now describe the effect of retinoids on the binding of EGF to its receptor. Retinoic acid enhances binding of ¹²⁵I-labelled EGF to various fibroblastic and epidermal cell lines. It has no marked effect on the affinity of this growth factor for its receptor, but increases the number of EGF receptor sites. Retinoic acid has little effect on the binding of concanavalin A (Con A) and insulin, indicating the specific nature of the action of retinoids on cell-surface glycoproteins. Treatment of cells with the phorbol ester 12-*O*-tetradecanoyl phorbol-13-acetate (TPA) and retinoic acid shows poor antagonism between these compounds on EGF binding. It has been previously shown that retinoids induce or stimulate differentiation of embryonal carcinoma cells^{11–13}. EGF binding can be used as a marker to monitor differentiation of these cells.

When sparse cultures of 3T6 cells are treated with retinoic acid the growth rate and saturation density are reduced, cells become more adhesive and have less tendency to overlap⁶. Figure 1 shows the effect of various concentrations of retinoic acid and its phenyl analogue on the binding of ¹²⁵I-labelled mouse EGF to 3T6 cells. Retinoic acid markedly increases binding of ¹²⁵I-EGF to its receptor, whereas the phenyl analogue does not affect this binding. Maximum binding, five- to sevenfold above control levels, occurs at a concentration of

3×10^{-6} M retinoic acid. The effect of retinoic acid on EGF binding is reversible; 3 days after the removal of retinoic acid, levels of EGF binding are reduced to those of control cells (not shown).

Because for certain cell lines, EGF binding is strongly dependent on cell density^{15,16}, the retinoic acid-induced enhancement of EGF binding might be related to the reduction of saturation density caused by this compound. However, we found no dependency of EGF binding on cell density in either retinoic acid-treated 3T6 cells ($1-6 \times 10^4$ cells cm^{-2}) or control cells ($1-12 \times 10^4$ cells cm^{-2}).

To determine the structural specificity of retinoids for stimulating EGF binding, we studied the effect of several analogues on the binding of this growth factor to 3T6 cells (Table 1). Retinoic acid and its *cis* analogue are the most potent derivatives tested, whereas the phenyl, pyrimidyl, methyl ether and ethylamide analogues are inactive. Retinol and retinyl acetate also increase EGF binding, although less effectively than retinoic acid. The specificity of the various analogues for stimulating EGF binding parallels that with which they increase adhesiveness and reduce saturation density of 3T6 cells² and induce differentiation of embryonal carcinoma cells¹³. Furthermore, except for retinol and retinyl acetate, which do not bind to the retinoic acid-binding protein, the effect of all analogues on EGF binding correlates exactly with their capacity to compete with retinoic acid for the binding site of this binding protein^{2,13,17}. We have found that retinyl acetate is rapidly converted to retinol and that 3T6 cells do not contain detectable amounts of retinol-binding protein (unpublished results). The activity of both retinol and retinyl acetate, therefore, cannot be due to the involvement of this binding protein but may be explained by the conversion of retinol to a metabolite that is able to bind effectively to the retinoic acid-binding protein.

To investigate whether the enhancement of EGF binding induced by retinoids is due to either an increase in unoccupied receptors or increased affinity, the effect of EGF concentration was determined. As shown in Fig. 2, retinoic acid has little effect on the affinity of ¹²⁵I-labelled mouse EGF for its receptors. Scatchard plot analyses indicate dissociation constants of 4.1×10^{-10} M and 4.5×10^{-10} M for retinoic acid-treated and control cells, respectively. However, the number of unoccupied EGF receptor sites is markedly increased from 2.87×10^4 receptors per cell for control cells to 1.51×10^5 receptors per cell after retinoic acid treatment.

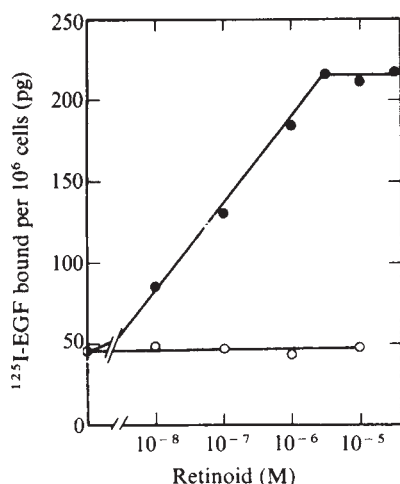


Fig. 1 Effects of retinoic acid at various concentrations on the binding of ^{125}I -labelled mouse EGF to 3T6 cells. Cells were grown in DMEM containing 10% calf serum for 3 days in the presence of either retinoic acid dissolved in ethanol or ethanol alone (0.15% v/v final concentration). After trypsinisation cells were subcultured in 35-mm cluster dishes ($1-2 \times 10^5$ cells per well). After 48 h incubation, binding of EGF was determined in duplicate as described by Todaro *et al.*^{16,18}. The medium was aspirated and the cells washed twice with 2 ml binding buffer (DMEM containing 1 mg ml^{-1} bovine serum albumin and 50 mM N,N -bis-(2-hydroxyethyl)-2-aminoethanesulphonic acid adjusted to pH 6.8). In the binding assays 1 ml binding buffer containing 0.825 ng of ^{125}I -labelled mouse EGF (Collaborative Research, 1.4×10^5 d.p.m.) was added to each well. After 60 min at 37°C unbound ^{125}I -EGF was removed, cells were washed four times with 2 ml binding buffer and solubilised in 0.75 ml lysing buffer (0.1 M Tris-HCl, pH 7.4, containing 0.5% SDS and 1 mM EDTA) and the wells washed twice with 0.5 ml lysing buffer. Radioactivity was counted in a Packard γ -counter. Nonspecific binding was determined in the presence of $10 \mu\text{g}$ unlabelled mouse EGF (Collaborative Research) and consisted of 100–300 d.p.m. per 10^6 cells. Cell number was determined in a Coulter counter; in all cases binding was determined at equal cell density ($\sim 5 \times 10^5$ cells per well). ●, Retinoic acid; ○, phenyl analogue of retinoic acid.

To test the generality of its effect on EGF binding, we studied the action of retinoic acid on several other cell lines and two other ligands, Con A and insulin (Table 2). Retinoic acid increases the binding of EGF in several other mouse fibroblast cell lines, including two clones of 3T12, two of 3T3 and one 3T3 derivative obtained by transformation with the chemical carcinogen dimethylbenzanthracene. Another 3T3 derivative transformed by benzo(a)pyrene shows very low levels of EGF binding which are not increased by treatment with retinoic acid. Also, melanoma S91M3 cells show very low binding of EGF, whereas treatment with retinoic acid enhances this binding. Retinoic acid stimulates severalfold the binding of EGF to two mouse epidermal cell lines, JB-6 and its tumorigenic derivative T62 obtained after selection by treatment with TPA¹⁸ but has little effect on EGF binding to human epidermoid carcinoma cells A431 and to Syrian hamster kidney cells BHK-21. Treatment of 3T3 cells with retinoic acid inhibits cell growth, enhances cell-to-cell substratum adhesiveness, increases production of glycosaminoglycans and stimulates EGF binding, whereas none of these characteristics is affected in its virus-transformed derivative 3T3SV. This difference can be attributed, respectively, to the presence and absence of retinoic acid-binding protein in these two cell lines².

Retinoic acid treatment has little effect on Con A and insulin binding (Table 2). It increases Con A binding to 3T6 and 3T12 C1A cells by 41% and 29%, respectively, but this binding is virtually unchanged in 3T3 and 3TSV cells. Retinoic acid does not affect binding of insulin to any of these cell lines. These results indicate a specific nature of the action of retinoids on cell-surface glycoproteins.

Table 1 The effect of retinoids on the binding of ^{125}I -labelled mouse EGF to 3T6 cells

Basic chemical structure	R	^{125}I -EGF binding (pg per 10^6 cells)	Relative binding
	CH_2OH^*	230.4	2.71
	$\text{CH}_2\text{OCOCH}_3$	257.6	3.03
	$\text{COOH}^\dagger$	483.2	5.68
	CH_2OCH_3	97.0	1.14
	COOH	324.3	3.81
	CONHC_2H_5	91.2	1.07
	COOH	478.2	5.63
	COOH	83.2	0.98
	COOH	356.0	4.18
	COOH	71.6	0.84

Cells were grown for 5 days in the presence of 10^{-5} M of the indicated retinoid. The binding of ^{125}I -EGF was determined as described in Fig. 1 legend; $1.25 \text{ ng } ^{125}\text{I}$ -EGF per ml (counts 1.2×10^5 d.p.m.) was added. Control cells bound 85.0 pg per 10^6 cells. The relative binding is defined as binding of ^{125}I -EGF to retinoid-treated cells/binding to control cells.

^{*}, Retinol; [†], retinoic acid.

Several reports have shown antagonism between retinoic acid and the tumour promoters phorbol esters. Retinoic acid inhibits induction of ornithine decarboxylase by phorbol esters¹⁹ and prevents growth in soft agar induced by sarcoma growth factor and phorbol esters^{20,21}. Phorbol esters reduce EGF binding by decreasing its affinity for its receptor¹⁴. When TPA is added to 3T6 or 3T12 C1A cells, EGF binding is reduced within 5 h, whereas EGF binding to retinoic acid-treated cells increases gradually with time, becoming optimal after 3 days of treatment (Table 3). Cells incubated simultaneously for 5 h with TPA and retinoic acid show no antagonism between these compounds on

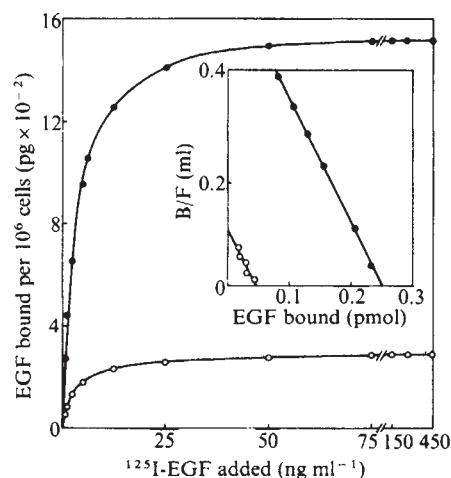


Fig. 2 Effect of EGF concentration on binding of ^{125}I -labelled mouse EGF to retinoic acid-treated and control 3T6 cells. Cells were treated with 10^{-5} M retinoic acid for 5 days. The binding assays were carried out as described in Fig. 1 legend. Data in the main figure were used for Scatchard plot analyses (inset). ○, Control; ●, retinoic acid treated cells; B/F, bound/free EGF.

Table 2 Effect of retinoic acid on the binding of EGF, Con A and insulin to various cell lines

Cell type	Designation	Binding of ^{125}I -ligand (10^3 d.p.m. per 10^6 cells)					
		EGF		Con A		Insulin	
		Control	RA-treated	Control	RA-treated	Control	RA-treated
Mouse fibroblast	3T6	3.82	25.98 (6.8)	48.40	68.24 (1.41)	1.06	0.89 (0.84)
Mouse fibroblast	3T12CLA	2.18	16.54 (7.6)	32.36	41.83 (1.29)	0.61	0.60 (0.98)
Mouse fibroblast	BALB/3T3	4.59	15.14 (3.3)	32.66	35.10 (1.07)	0.54	0.53 (0.98)
Mouse fibroblast	BALB/3T3SV	4.92	5.93 (1.2)	22.10	21.70 (0.98)	0.57	0.57 (1.0)
Mouse fibroblast	3T12	4.16	9.66 (2.3)				
Mouse fibroblast	BALB/3T3A31-1-1*	1.74	6.20 (3.56)				
Mouse fibroblast	BALB/3T3A31-1-BP-2*	0.14	0.16 (1.14)				
Mouse fibroblast	BALB/3T3A31-1-DMBA-1*	2.68	11.42 (4.26)				
Mouse epidermal	JB-6†	3.70	9.36 (2.53)				
	T62‡	1.74	5.90 (3.4)				
Mouse melanoma	Cloudman S-91 M3	0.03	0.20 (6.7)				
Syrian hamster	BHK-21	3.81	3.04 (0.8)				
kidney							
Human epidermoid	A431‡	51.12	49.44 (0.97)				
carcinoma							

The binding assays were carried out in duplicate as described in Fig. 1 legend. All cells were grown in DMEM containing 10% calf serum except the melanoma S91M3 cells, which were grown in Dulbecco's modified Eagle's medium (DMEM) containing 15% horse serum and 2.5% fetal calf serum. All cells were treated for 5 days with 10^{-5} M retinoic acid (RA) or with ethanol in the control and were at equal density (5×10^5 cells per 35-mm dish) at the time of the binding assay. The specificities of the ligands used were: ^{125}I -labelled mouse EGF 0.825 ng ml^{-1} (9.5×10^4 d.p.m. per ng) was added, with $10 \mu\text{g}$ of unlabelled EGF to correct for nonspecific binding; ^{125}I -labelled Con A (NEN) 25 ng ml^{-1} (9.8×10^5 d.p.m.) and 15 mg ml^{-1} of α -methyl-D-mannoside to correct for nonspecific binding; ^{125}I -labelled insulin (Sigma) 2.5 ng ml^{-1} (2.7×10^5 d.p.m.) and $20 \mu\text{g ml}^{-1}$ unlabelled insulin to correct for nonspecific binding. The relative binding (binding of ^{125}I -labelled ligand to retinoic acid-treated cells/binding to control cells) is given in parentheses.

*Obtained from Dr T. Kakunaga; †obtained from Dr. N. Colburn; ‡obtained from Dr G. J. Todaro.

Table 3 Comparison of the action of retinoic acid and TPA on the binding of EGF to 3T6 and 3T12 C1A cells

Cell line	Treatment	^{125}I -EGF binding (pg per 10^6 cells)	Relative binding
3T6 (grown in the presence of 0.15% ethanol)	Control	87.0	1.0
	+TPA, 5 h	4.8	0.06
3T6 (grown in the presence of 10^{-5} M RA for 5 d)	Control	443.9	5.10
	+TPA, 5 h	77.7	0.89
3T12 CLA (grown in the presence of 0.15% ethanol)	Control	34.1	1.0
	+TPA, 5h	5.4	0.16
	+RA, 5 h	48.4	1.42
	+RA, 24 h	93.5	2.74
	+RA, 48 h	131.7	3.86
	+TPA, +RA, 5 h	5.6	0.16
3T12 CLA (grown in the presence of 10^{-5} M RA for 5 d)	Control	175.6	5.15
	+TPA, 5 h	86.5	2.54

Cells were treated with 100 ng ml^{-1} TPA and with 10^{-5} M retinoic acid (RA) at the times indicated. Controls received identical amounts of dimethyl sulphoxide or ethanol. EGF binding was assayed as described in Fig. 1 legend, EGF concentration was 1.25 ng ml^{-1} (1.1×10^5 d.p.m. per ng).

the binding of EGF. Cells grown for 5 days in the presence of retinoic acid also exhibit a reduced binding on TPA treatment, although this decrease is relatively smaller than in control cells. Thus, retinoic acid poorly antagonises the effect of TPA on EGF binding. Similar results were obtained for two epidermal cell lines, JB-6 and T62 (L.D. Dion and A.M.J., unpublished results). The above results seem to indicate an essential difference between the mode of action of retinoids and phorbol esters on EGF receptors: the slow action of retinoids on the cell surface seems to involve metabolism and/or macromolecular syntheses whereas the rapid action of phorbol esters may reflect conformational changes induced by alterations in the fluid state of the membrane²².

We have shown recently that retinoic acid stimulates differentiation of embryonal carcinoma cells PCC4 aza 1R (refs 12, 13). During 24–48 h of treatment with retinoic acid these

cells differentiate irreversibly into a cell type with a fibroblastic morphology. The differentiated cells can be distinguished from the original embryonal carcinoma cells by a variety of cell-surface markers¹². The PCC4 aza 1R cells exhibit very low levels of EGF binding whereas the differentiated derivative PCC4 D shows much higher levels of binding (Table 4). Therefore, EGF binding can be used as a marker to monitor differentiation of PCC4 aza 1R cells. After 24 h of treatment with retinoic acid, EGF binding increases concomitantly (not shown) with the appearance of cells with a fibroblastic morphology and the production of plasminogen activator which was used previously to monitor differentiation of PCC4 aza 1R cells¹². Similar results were recently reported for other embryonal carcinoma cell lines²³. The phenyl analogue of retinoic acid, which is unable to stimulate differentiation of embryonal carcinoma cells, has no stimulatory effect on EGF binding.

Thus, retinoic acid treatment of several cell lines enhances the binding of mouse EGF to its receptors, not by changing its affinity but by increasing the number of available receptor sites. Stimulation of EGF binding may result from an increase either in total receptor sites or in the number of unoccupied sites alone. Certain cell lines are known to produce endogenous growth factors²⁰, some of which may bind to the EGF receptor;

Table 4 Increase in EGF binding during differentiation of embryonal carcinoma cells PCC4 aza 1R induced by retinoic acid

Cell line	Treatment	^{125}I -EGF binding (pg per 10^6 cells)
PCC4 aza 1R	Control	1.2
	24 h RA	1.4
	48 h RA	20.3
	72 h RA	58.9
	96 h RA	112.7
	96 h ethanol	1.5
	96 h phenyl analogue	1.3
PCC4 D	Control	67.1
	96 h RA (10^{-5} M)	143.4

Cells were grown in DMEM containing 10% fetal calf serum. Differentiation was induced by the addition of 10^{-6} M RA or ethanol in the control. EGF binding was determined at the time intervals indicated, 1 ng EGF per ml (1.2×10^5 d.p.m.) was used. PCC4 D is the differentiated derivative obtained from PCC4 aza 1R after cloning¹¹.

retinoids may suppress the production of these factors leading to an increase in the number of unoccupied receptors and decreased cell growth. Alternatively, retinoids may increase the synthesis of EGF receptors. The implications of the increase in EGF receptor for cell growth remain to be established. Preliminary results show that EGF can enhance mitotic activity in serum-starved retinoic acid-treated cells, indicating that retinoids do not totally block the EGF response²⁶.

The action of retinoids on EGF binding provides us with a tool for studying the involvement of retinoid metabolites, like mannosylretinyl phosphate^{1,7}, in glycosylation of glycoproteins and a fast screening test for retinoids potentially active in inhibiting carcinogenesis^{24,25}.

I thank Dr Luigi M. De Luca, in whose laboratory this work was carried out, for his support and comments on the manuscript, Dr L. D. Dion for his comments on the manuscript and Dr Beverly A. Pawson (Roche) for providing the retinoid analogues.

Received 8 November 1979; accepted 13 February 1980.

- De Luca, L. M. *Vitam. Horm.* **35**, 1–57 (1977).
- Jetten, A. M., Jetten, M. E. R., Shapiro, S. & Poon, J. *Expl Cell Res.* **119**, 289–299 (1979).
- Hassell, J., Pennypacker, J., Kleinman, K., Pratt, R. & Yamada, K. *Cell* **17**, 821–826 (1979).
- Patt, L., Itaya, K. & Hakomori, S. *Nature* **273**, 379–381 (1978).
- Adamo, S., De Luca, L. M., Akalovsky, I. & Bhat, P. J. *natn. Cancer Inst.* **62**, 1473–1478 (1979).
- Lotan, R. *Biochim. biophys. Acta* (in the press).
- De Luca, L. M. *et al. Pure appl. Chem.* **51**, 581–591 (1979).
- Todaro, G. J., De Larco, J. & Cohen, S. *Nature* **264**, 26–31 (1976).
- Carpenter, G. & Cohen, S. *A. Rev. Biochem.* **48**, 193–216 (1979).
- Gospodarowicz, D., Moran, J. & Mescher, A. in *Molecular Control of Proliferation and Differentiation* (eds Papaconstaniou, J. & Rutter, W. T.) 32–63 (Academic, New York, 1978).
- Strickland, S. & Mahdavi, V. *Cell* **15**, 393–404 (1978).
- Jetten, A. M., Jetten, M. E. R. & Sherman, M. *Expl Cell Res.* **124**, 381–391 (1979).
- Jetten, A. M. & Jetten, M. E. R. *Nature* **278**, 180–182 (1979).
- Shoyab, M., De Larco, J. & Todaro, G. J. *Nature* **279**, 387–391 (1979).
- Pratt, R. & Pastan, I. *Nature* **272**, 68–70 (1978).
- Brown, K., Yeh, Y. & Holley, R. W. *J. cell. Physiol.* **100**, 227–238 (1979).
- Chytil, F. & Ong, D. in *Receptors and Hormone Action* Vol. II (eds O'Malley, B. W. & Birnbaumer, L.) 573–591 (Academic, New York, 1978).
- Colburn, N., Former, B., Nelson, K. & Yuspa, S. *Nature* **282**, 589–591 (1979).
- Boutwell, R. & Verma, A. *Pure appl. Chem.* **51**, 857–886 (1979).
- Todaro, G. J., De Larco, J. & Sporn, M. *Nature* **276**, 272–274 (1978).
- Colburn, N. in *Neoplastic Transformation in Differentiated Epithelial Cell Systems In Vitro* (eds Franks, L. & Wigley, C.) 113–134 (Academic, New York, 1979).
- Fisher, P., Flamm, M., Schachter, D. & Weinstein, I. *Biochem. biophys. Res. Commun.* **86**, 1063–1068 (1979).
- Rees, A., Adamson, E. & Graham, C. *Nature* **281**, 309–311 (1979).
- Sporn, M., Dunlop, N., Newton, D. & Smith, J. *Fedn Proc.* **35**, 1332–1338 (1976).
- Bollag, W. *Eur. J. Cancer* **8**, 689–693 (1972).
- Jetten, A. M. *Ann. N.Y. Acad. Sci.* (in the press).

Vitamin C preferential toxicity for malignant melanoma cells

Stanley Bram*†, Patrick Froussard*, Marcelle Guichard†, Claude Jasmin*, Yvette Augery*, Françoise Sinoussi-Barre‡ & Wayne Wray§

* Institute of Cancerology and Immunogenetics, Hôpital Paul Brousse, 94800 Villejuif, France

† Institut Gustave Roussy, 94800 Villejuif, France

‡ Institut Pasteur, 75015 Paris, France

§ Baylor University Medical School, Houston, Texas 77030

Vitamin C has been suggested^{1,2} and disputed³ as an anti-cancer agent. For cells in culture, no preferential effect against any type of cancer has yet been demonstrated. Our aim here is to show that vitamin C is selectively toxic to at least one type of malignant cell—a melanoma—at concentrations that might be attained in humans. Copper ions react with ascorbate and generate free radicals in solution⁴. Ascorbate when combined with copper rapidly reduces the viscosity of DNA solutions and has exhibited some carcinostatic effects on transplanted sarcoma 180 tumours in mice⁵. We reasoned that the elevated copper concentration in melanoma^{6,7} could result in a more selective toxicity for ascorbate.

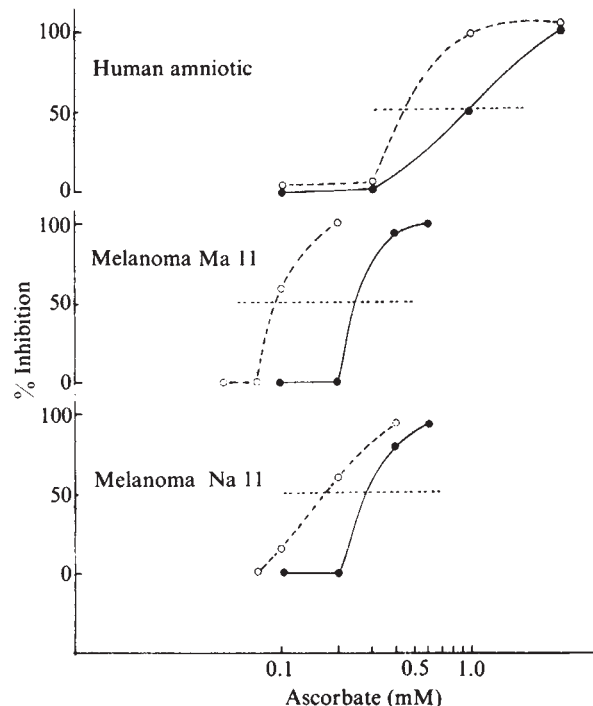


Fig. 1 A typical experiment on two melanoma and one normal cell line determining the reduction in the number of colonies by ascorbate only (●); or ascorbate + 5 μM copper (○). Colony formation, which directly measures reproductive death, was carried out as previously described⁸. Briefly, about 200–500 single cells were inoculated into culture boxes and the number of colonies containing more than 50 cells was counted 8–14 days later. Colony-forming efficiency of parallel controls ranged from 25 to 60% for the different cell lines studied with this technique.

Inhibition was obtained from:

$$\frac{\text{No. of untreated control colonies} - \text{No. of treated colonies}}{\text{No. of untreated control colonies}}$$

Each point represents the average of a triplicate determination.

Cells (see Table 1) were grown in McCoy's medium supplemented with 15% fetal calf serum and antibiotics. Cytotoxic effects were determined using three independent techniques with parallel simultaneous runs on a melanoma and a non-melanotic cell culture (see Fig. 1). Drugs were added to cells 16–24 h after inoculation with single cell suspensions. Sodium ascorbate (Sigma) was added to cells from a water stock solution within 2 h of preparation.

Table 1 summarises the results obtained on 13 cell lines in four different laboratories. The concentration of ascorbate which resulted in a 50% decrease in colony formation, cell number or viability was between 0.2 and 0.5 mM (a mean of 0.3) for all the melanotic cell cultures studied. Cloned cell lines with many trypsinisations were as sensitive as cells recently put into culture. No obvious difference relating to the degree of pigmentation was noted. Ascorbate was added to Na11 cells which were suspended in agar⁸ 30 min after trypsinisation, and in another instance 30 h after inoculating into a Falcon box, yet the differences in the ID₅₀ were not significant. Plots of percentage inhibition with the three measuring techniques used not only gave similar ID₅₀ concentrations, but also exhibited a very sharp increase in inhibition at about this value (Fig. 1).

The non-melanoma cells tested (Table 1) were less sensitive to ascorbate and these other cell types exhibited a much larger range in their ID₅₀ values. Only mouse embryo fibroblast cultures showed an ascorbate sensitivity approaching that of melanoma cells.

A catalytic concentration of copper (5 μM or ~20 less than the Cu²⁺ ID₅₀) increased ascorbate toxicity for melanoma two- to fivefold (Table 1). Cytotoxicity to other cells increased only 0 to 1.5 times.

Table 1 Effect of ascorbate on the growth of melanoma and other cells

	Cells			ID ₅₀ * (mM)			
	Pigmentation	Doubling time (h)	No. of passage	Measuring technique	Ascorbate	Ascorbate + 5 μ M Cu ²⁺	Cu ²⁺
<i>Melanoma (human)</i>							
FO1 ^a	+	39	100	t.b. ²	0.2	0.05	0.2
Bel ^a	+	30	105	g.i. ¹	0.4	—	—
Val ^a	±	30	20	g.i. ¹	0.4	—	—
Na11 ^b	++	20	160	c.f. ³	0.3	0.1	0.2
				c.f.a. ¹	0.15	—	—
				g.i. ³	0.3	0.07	0.2
Ma11 ^b	+	22	80	c.f. ³	0.25	0.09	0.1
DOR ^c	0	48	100	g.i. ¹	0.4	—	0.25
Mouse B16	+	48	3	c.f. ²	0.5	0.05	0.01
				Average	0.3	0.07	0.15
<i>Non-melanoma</i>							
Normal human amniotic cells ^a		24	60	c.f. ²	1.0	0.6	0.1
				g.i. ³	1.0	0.6	0.2
Human fibroblasts (F1) ^d		24	3	g.i. ¹	0.9	0.9	0.4
Mouse embryo fibroblasts ^f		20	30	c.f. ¹	0.4	0.4	0.2
				g.i. ¹	0.7	0.7	0.2
<i>Transformed</i>							
Chinese hamster ovary		12	>100	t.b. ²	2.0	1.5	0.5
Chinese hamster V79 ^b		16	>100	c.f. ¹	0.7	0.5	0.3
HeLa (human)		24	>100	t.b. ²	1.0	0.8	0.5
				Average	1.0	0.7	0.3

g.i., Growth inhibition measurements were carried out by inoculating 10⁵ cells into 25-cm² boxes (Falcon). Drug exposures from 2 to 5 days with one to four ascorbate additions were studied. Results shown were obtained by addition of ascorbate 1 day after inoculation and by trypsinising and counting cells in a haemocytometer on day 5. Per cent inhibition was calculated from the equation in Fig. 1 using the number of cells counted.

c.f., Colony formation as described in Fig. 1.

c.f.a., colony formation in agar was carried out as by Salmon *et al.*⁸

t.b., Trypan blue exclusion (0.2%) and cell morphology were determined 5 days after one drug addition to about 1,000 cells in Falcon microwell plates. Viability determinations were done each in triplicate.

* Average value calculated from curves like those in Fig. 1.

^{1,2,3}, The number of independent experiments, each performed in duplicate or triplicate.

a, Cells from Dr B. Giovannella; b, see ref. 9 for further details; c, cells from Dr N. Saal; d, cells put into culture in September 1979 from normal adult skin; e, cells from Flow Laboratories; f, C₃H cells from Dr R. H. Bassin.

The results can be described as follows: melanoma cells were 2–10-fold and 5–20-fold, respectively, more sensitive to ascorbate, and to ascorbate plus 5 μ M Cu²⁺ than for amelanotic cells. Referring to the typical inhibition curves of Fig. 1, it can be seen that in 0.6 mM ascorbate the ability of melanoma cells to form colonies is 10 to 20 times less than for normal human amniotic cells. In 1 mM ascorbate no melanoma colonies were observed, while ~300 colonies were counted in controls; yet, at this concentration amniotic cell colony formation was inhibited by only 50%. The ratio of relative inhibition is then ~100-fold here.

With the exception of melanoma, no obvious difference between transformed (malignant) cell lines and normal cells was observed (Table 1). Copper ions alone exhibited some preferential toxicity to melanoma cells. On the other hand, the sensitivity to Cu²⁺ varied greatly from one cell line to another (Table 1), and non-transformed human amniotic cells were killed by lower concentrations of copper than many pigment cells.

The inhibitory effect of copper as sulphate, gluconate and acetate salts was studied. Although the pH of the stock 0.1 M solution added varied from 3.5 with copper sulphate to 5.3 with copper acetate, no difference in the Cu²⁺ ID₅₀ for melanoma cells was observed. Ceruloplasmin, the usual serum transport protein for copper, was not toxic to either melanoma or amniotic cells in the range 10–100 μ g ml⁻¹. However, in a preliminary study, both 7 × 10⁻⁵ M ascorbate and 30 μ g ml⁻¹ of ceruloplasmin caused a 50% inhibition in colony formation.

Note that the relative growth inhibition of vitamin C for melanoma compared to normal cells at 5 × 10⁻⁴ M (Table 1) is similar to that reported for vitamin A at 10⁻⁵ M (ref. 10). However, vitamin C is about 1,000 times less toxic than vitamin A.

The only previous report of ascorbate toxicity on melanoma cells was part of a study of the killing effect of dopa and other melanin precursors¹¹. Our results confirm the finding¹¹ that ascorbate reduces the toxic effect of dopa. The pertinent result mentioned in ref. 11 was that 1 mM ascorbate reduced thymidine incorporation in Cloudman S-91 cells by 33% after a

16-h exposure. We obtained a similar inhibition with a 16-h exposure. In our experiments we have found that cytotoxic effects at 2–4 × 10⁻⁴ M ID₅₀ concentrations of ascorbate become evident only after about 48 h. For non-melanoma cells, our ID₅₀ of 1–2 mM for human amniotic, HeLa, and CHO cultures compares well with values in the literature (Table 2). Vitamin C and copper are required for melanin biosynthesis, and are preferentially incorporated by melanoma tissues^{6,7,16}, yet, at elevated concentrations, both of these agents inhibit dopa oxidase activity^{17,18}. This is one possible mechanism for the specific inhibition of melanoma cell proliferation.

Nonspecific ascorbate toxicity has been attributed to the formation of H₂O₂ during its auto-oxidation^{14,19}, but at pH < 7.6 ascorbate is no longer auto-oxidisable except in the presence of metallic catalysts such as copper¹³. Ascorbate has little or no antimicrobial activity in the absence of metal ions⁴. A cogent example from the early literature is that 6 × 10⁻⁵ M ascorbate plus 5 μ M Cu²⁺ completely inactivated polio virus after a 4-h contact⁴, while in the presence of EDTA, 6 × 10⁻⁴ M ascorbate had no effect. Thus it seems reasonable to suggest that the selective toxicity of ascorbate for melanoma is partly due to preferential incorporation of copper. That catalytic concentrations of Cu²⁺ greatly increase the preferential toxicity of ascorbate for melanoma cells, suggests that an intracellular mechanism is involved here.

A likely target for ascorbate killing is DNA. Mutagenic properties of ascorbate and copper have been demonstrated on

Table 2 Ascorbate inhibition

Cell types	Approximate ID ₅₀ (mM)	Refs
Rat plasmacytoma	1.5*	12
Mouse neuroblastoma	1.5*	13
Mouse glioma	2*	13
Ehrlich ascites	10†	14
Chinese hamster V 79	4†	15

* An exposure to ascorbate for more than 36 h was used as in Table 1.

† An exposure of less than 4 h was used.

bacteria and bacteriophage^{21,22} and more recently on eukaryotic cells²³. Vitamin C and copper rapidly eliminate the replication and virulence of bacteria (that is, *Bacillus pestis* of bubonic plague), but do not alter neither morphology antigenicity, or most biochemical reactions over a much longer period of time²¹.

Ascorbate rarely acts as a coenzyme or vitamin but generally modifies the redox potential of the cell. Vitamin C can thereby appreciably potentiate or deactivate many antitumour drugs¹³.

The results presented here indicate that vitamin C may directly inhibit the growth of proliferating cells, and this might explain some of the reported carcinostatic effects². Such an inhibition would depend on the extent to which a given cell incorporates ascorbate, and the sensitivity of the cell to its

cytotoxic effects. Melanoma cells have been found to preferentially incorporate vitamin C (ref. 16), and our *in vitro* studies show that vitamin C is more toxic to melanotic cells than to any others studied so far.

Saturation levels of ascorbate in tissues range from 1 to 2 mM (ref. 24). Such concentrations in our experiments inhibit malignant melanoma cells 20 to 500 times more than the other cell studied. Moreover, as 10^{-5} M free copper concentrations are not toxic in man, the synergistic effect of copper on ascorbate can also be exploited. *In vivo* experiments on mice with melanoma are under way, and clinical studies are planned.

We thank Drs E. Malaise and J. C. Chermann for suggestions and Dr B. Giovanella for his gift of cell lines. Financial support was provided by the CNRS, INSERM and the DGRST.

Received 7 December 1979; accepted 11 February 1980.

1. Deucher, W. G. *Strahlentherapie* **67**, 143–151 (1940).
2. Cameron, E., Pauling, L. & Libovitz, B. *Cancer Res.* **39**, 663–681 (1979).
3. Creagan, E. T. *et al.* *New Engl. J. Med.* **301**, 687–690 (1979).
4. Peloux, Y., Nofre, C., Cier, A. & Colbert, L. *Annls Inst. Pasteur, Paris* **102**, 6–23 (1962).
5. Yamafuji, K., Nakamura, Y., Omura, H., Soeda, T. & Gyotoku, K. *Z. Krebsforsch.* **76**, 1–7 (1971).
6. Ikonopisov, R. L. in *Melanoma and Skin Cancer* (ed. McCarthy, W. H.) 223–230 (International Union against Cancer, Sidney, 1972).
7. Horcicko, J., Borovansky, J., Duchon, B. & Prochazkova, B. *Sb. lék.* **79**, 340–343 (1977).
8. Salmon, S. *et al.* *New Engl. J. Med.* **298**, 1322–1327 (1978).
9. Weininger, J., Guichard, M., Joly, A. M., Malaise, E. P. & Lachet, B. *Int. J. Radiat. Biol.* **34**, 285–290 (1978).
10. Lotan, R., Giotta, G., Nork, E. & Nicolson, G. L. *J. natn. Cancer Inst.* **60**, 1035–1041 (1978).

11. Benade, L., Howard, T. & Burk, D. *Oncology* **23**, 33–43 (1969).
12. Jungeblut, C. W. *J. exp. Med.* **62**, 517–521 (1935).
13. Barron, E., Meio, R. & Klemphofer, F. *J. biol. Chem.* **112**, 625–632 (1936).
14. Dodin, A. *C. r. hebd. Séanc. Acad. Sci., Paris* **258**, 2707–2710 (1964).
15. Lwoff, A. *Annls Inst. Pasteur, Paris* **84**, 225–241 (1953).
16. Stich, H. F., Karim, J., Koropatnick, J. & Lo, L. *Nature* **260**, 722–723 (1976).
17. Prasad, K. N., Sinha, P. K., Ramanujam, M. & Sakamoto, A. *Proc. natn. Acad. Sci., U.S.A.* **76**, 829–832 (1979).
18. Dyer, H. M. & Ross, H. E. *J. natn. Cancer Inst.* **11**, 313–318 (1950).
19. Chatterjee, G. C. in *The Vitamins* (eds Sebel, L. W. & Harris, R. S.) 399–406 (Academic, New York, 1969).
20. Park, C. H., Bergsagel, D. E. & McCalloch, E. A. *Science* **174**, 720–721 (1971).
21. Koch, C. J. & Biaglow, J. E. *J. cell. Physiol.* **94**, 299–306 (1978).
22. Pawelek, J. M. & Lerner, A. B. *Nature* **276**, 627–628 (1978).
23. Tomita, Y. & Seiji, M. *J. Derm.* **4**, 245–249 (1977).
24. Cooper, M. & Maclean, N. *Yale J. Biol. Med.* **46**, 705–706 (1973).

Vasopressin increases water permeability by inducing pores

Stephen Gluck & Qais Al-Awqati

Departments of Medicine and Physiology, Columbia University, College of Physicians and Surgeons New York, New York 10032

Vasopressin is thought to increase the water permeability of the luminal membrane of distal urinary epithelia either by enhancing the solubility or diffusion of water in the lipid phase¹ or by inducing aqueous pores². Measurement of the permeability to water and small solutes in different experimental conditions should allow the distinction between these two modes of water transport^{1–9}, but because of technical difficulties (for instance the effect of unstirred layers^{3–6}) or conceptual problems (in the case of activation energy measurements^{7–9}) the interpretation of many of these methods is in question. We present here a new approach to this problem using proton conductance of the membrane as a probe for the presence of aqueous channels. Protons travel in free solution by jumping from one water molecule to another¹⁰. If vasopressin induces aqueous pores, it should also increase proton permeability, as protons will be able to jump from a water molecule in the bulk solution to another in the pore. However, if vasopressin changes water permeability by increasing membrane fluidity the proton conductance should not increase as the dielectric constant of even very fluid lipids is low, leading to extremely low solubility of ions in the lipid¹¹. We found that vasopressin increased the proton permeability of the luminal membrane of the toad urinary bladder and conclude that this hormone increases water permeability by inducing aqueous pores rather than by increasing lipid fluidity. Proton conductance may be of use in determining the existence and properties of other aqueous channels.

Vasopressin increased the proton disappearance from an acid mucosal medium bathing Dominican toad bladders, an epithelium which does not actively transport protons (Fig. 1). In 22 experiments vasopressin increased the proton permeability (G_H) by $7.24 \times 10^{-4} \text{ cm s}^{-1}$, a 300% increase.

To investigate the site of the vasopressin-induced proton conductance we measured the intracellular pH using the distribution of the weak acid, dimethyl-oxazolidine-dione

(DMO)¹². Paired bladders treated with vasopressin were placed in a solution containing ¹⁴C-DMO and ³H-inulin at pH 7.4 for 1 h. The mucosal solution of one member of each pair was replaced by a medium of pH 5.0 containing the same concentration of ³H-inulin and an amount of ¹⁴C-DMO calculated to give the same concentration of the free-acid DMO (~4% of the original DMO concentration). After 15 min, the

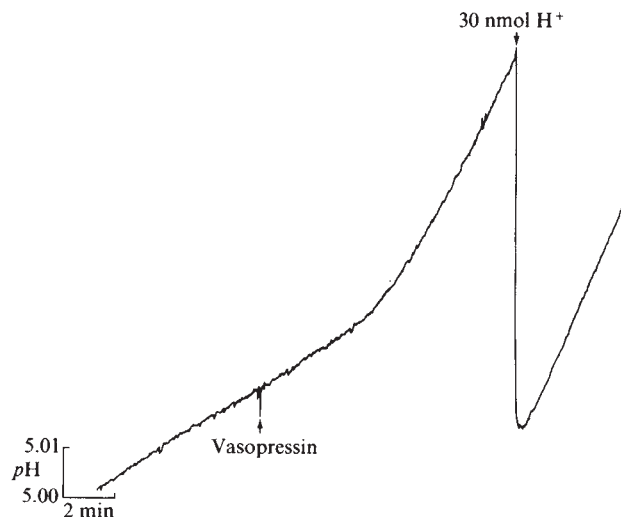


Fig. 1 Effect of vasopressin on H^+ movement across toad bladder. Bladders were placed in Ussing chambers, in short-circuit conditions, using a bicarbonate-free low buffer medium and rates of proton disappearance from the mucosal medium were measured by titration. The figure shows one experiment. Addition of vasopressin to the serosal medium markedly increases the rate of proton disappearance within a few minutes. When the pH reaches 5.1 a measured quantity of acid is added to the luminal medium serving to calibrate proton disappearance and to return the pH to 5. To demonstrate that the proton flux induced by vasopressin results from flow across the epithelium rather than from buffering at the surface, we showed in six experiments that the short circuit current in the presence of amiloride (which abolishes the sodium current) is not different from the simultaneously measured mucosal proton disappearance (0.42 compared with $0.36 \text{ nmol cm}^{-2} \text{ min}^{-1}$). Further, the mucosal proton disappearance and serosal proton appearance in the same bladders were equal (0.34 and $0.33 \text{ nmol cm}^{-2} \text{ min}^{-1}$, $n = 3$).

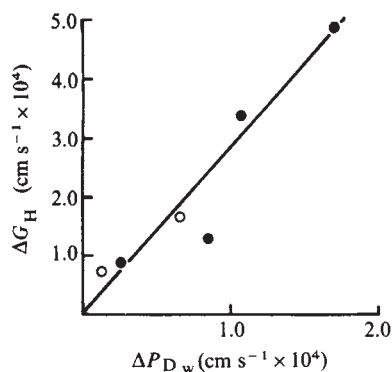


Fig. 2 Plot of the simultaneously measured changes in proton permeability (G_H) and $^3\text{H}_2\text{O}$ permeability (P_D) induced by vasopressin. The proton permeability was measured as the ratio of the flux to the concentration difference.

intracellular pH of scraped epithelial cells from the control bladder was 6.7 while the pH of those exposed to the low mucosal pH was 6.12 ($\Delta\text{pH} = 0.55 \pm 0.12$, $n = 6$). This suggests that the vasopressin-induced proton permeability occurs at the mucosal border of the cell. Further evidence for this was obtained from measurement of the effect of amiloride on H^+ fluxes in the presence of vasopressin. Amiloride increases the negative potential of the toad bladder epithelial cell and hence should increase the electrical gradient for proton movement into cells¹³ while having no effect on the paracellular conductance. In 10 experiments in vasopressin-treated toad bladders, addition of amiloride to the luminal medium increased the proton flux by $0.06 \pm 0.03 \text{ nmol cm}^{-2} \text{ min}^{-1}$ ($P < 0.05$). These results demonstrate that vasopressin increases the proton permeability of the luminal border of toad bladder cells, rather than the permeability of the paracellular pathway.

Simultaneous measurement of G_H and tritiated water permeability (P_D) showed that the increase in G_H correlated with the increase in P_D , $r = 0.947$, $n = 6$ (Fig. 2). Since vasopressin increases the urea and sodium permeability as well as water permeability, we investigated whether protons move through urea or sodium channels. Methohexital, an anaesthetic that inhibits vasopressin-induced water permeability but has little effect on urea and sodium permeabilities¹⁴, reduced the vasopressin-induced proton permeability from $5.80 \times 10^{-4} \text{ cm s}^{-1}$ to $1.44 \times 10^{-4} \text{ cm s}^{-1}$ ($P < 0.01$, $n = 9$). Two results suggest that the protons are not crossing urea channels. We found no correlation between ^{14}C -urea and proton permeability measured in 10 bladders. Furthermore, addition of phloretin, an agent which inhibits urea transport without affecting water permeability¹⁵, abolished the vasopressin-induced urea permeability (control $9.92 \times 10^{-4} \text{ cm s}^{-1}$, 0.1 mM phloretin $0.26 \times 10^{-4} \text{ cm s}^{-1}$, $P < 0.01$, $n = 6$ paired hemibladders) but had no significant effect on the proton permeability measured in the same bladders (control $2.60 \times 10^{-4} \text{ cm s}^{-1}$, phloretin $1.86 \times 10^{-4} \text{ cm s}^{-1}$). It is not likely that protons cross sodium channels as the effect of vasopressin on sodium transport was transient while its effect on proton permeability was not; 30 min after addition of vasopressin the proton flux was 561% above the initial rate while sodium transport was only 25% above the initial value ($n = 14$). Also, as shown above, amiloride abolished sodium transport but increased the proton flux in the presence of vasopressin.

Since protons, being small ions, cannot pass through even very fluid lipid membranes, the above results show that vasopressin must have induced aqueous channels which conduct protons. A rough estimate of the conductance of these channels can be obtained by comparing their proton to water permeability ratio to that of gramicidin A channels. The proton conductance of a single gramicidin A channel at pH 5.0 is 0.015 pS (ref. 16) and its P_D $1.8 \times 10^{-15} \text{ cm}^3 \text{ s}^{-1}$ (ref. 17), giving a ratio of 7.8 S s cm^{-3} . The vasopressin induced proton conductance was $3.22 \mu\text{S cm}^{-2}$ and the P_D increased by $0.78 \times 10^{-4} \text{ cm s}^{-1}$, giving a ratio of

0.04 S s cm^{-3} . The vasopressin-induced channels are at least 200 times less conductive than gramicidin A. Whether this is due to their native ion selectivity or to a very low density of water or to other factors remains to be determined.

Received 10 December 1979; accepted 13 February 1980.

- Pietras, R. J. & Wright, E. M. *Nature* **247**, 222–224 (1974).
- Finkelstein, A. J. *gen. Physiol.* **68**, 137–143 (1976).
- Koefoed-Johnsen, V. & Ussing, H. H. *Acta physiol. scand.* **28**, 60–76 (1953).
- Dainty, J. *Adv. Bot. Res.* **1**, 279–326 (1963).
- Hays, R. M. in *Curr. Topics Membranes Transport* **3**, 339–366 (1972).
- Schafer, J. A. & Andreoli, T. E. *J. clin. Invest.* **51**, 1264–1278 (1972).
- Leaf, A. & Hays, R. M. *J. gen. Physiol.* **45**, 933–948 (1962).
- Eggena, P. J. *gen. Physiol.* **59**, 519–533 (1972).
- Kachadorian, W. A., Muller, J., Rudich, S. W. & DiScala, V. A. *Science* **205**, 910–913 (1979).
- Eigen, M. *Angew. Chemie* **3**, 1–10 (1964).
- Parsegian, A. *Nature* **221**, 844–846 (1969).
- Waddell, W. J. & Butler, T. C. *J. clin. Invest.* **38**, 720–729 (1959).
- Sudou, K. & Hoshi, T. *J. Membrane Biol.* **32**, 115–132 (1977).
- Levine, S., Franki, N. & Hays, R. M. *J. clin. Invest.* **58**, 980–988 (1976).
- Levine, S., Franki, N. & Hays, R. M. *J. clin. Invest.* **52**, 1435–1442 (1973).
- Hladky, S. B. in *Drugs and Transport Processes* (ed. Callingham, B. A.) 193–210. (University Park Press, Baltimore, 1974).
- Rosenberg, P. A. & Finkelstein, A. J. *gen. Physiol.* **72**, 341–350 (1978).

A novel α -globin gene arrangement in man

D. R. Higgs, J. M. Old, L. Pressley, J. B. Clegg & D. J. Weatherall

MRC Molecular Haematology Unit, Nuffield Department of Clinical Medicine, John Radcliffe Hospital, Headington, Oxford OX3 9DU, UK

The human genome has two linked α -globin genes on chromosome 16. Deletion of one or more of them, as occurs in α -thalassaemia, leads to a reduced output of α -globin mRNA in proportion to the number of α -globin genes lost¹. In some racial groups deletion of one of the pair of α -globin genes may result from unequal crossing over between the genes on homologous chromosomes^{2,3} by a mechanism resembling that postulated for the formation of the $\delta\beta$ fusion genes of the Lepore haemoglobins⁴. By analogy, the opposite chromosome in this cross-over should have three α -globin genes just as the 'anti-Lepore' chromosome has three non- α -chain genes. We describe here a Welsh family in which three members have five α -globin genes—three on one chromosome and two on the other. The additional α gene results in an increased α mRNA output and it may therefore produce the phenotype of mild β -thalassaemia.

DNA was obtained from a 61-yr-old male (E.J.) and his sons (D.J., R.J. and C.J.). Two of the sons (D.J. and C.J.) were subsequently shown to have the same α -globin gene arrangement as their father whereas R.J. had a normal α -gene arrangement. Digestion of normal DNA with the endonucleases *Bam*HI and *Eco*RI, or a combination of the two, generates a single fragment containing both α -globin genes^{2,5,6}. When DNA from E.J. was digested with these enzymes two fragments were obtained, one of normal size and the other consistently 3.5–4.0 kilobases larger (Fig. 1A). This result suggested that E.J. has one chromosome which yields a normal restriction fragment and another which produces a larger fragment with an insertion of approximately 3.5–4.0 kilobases of DNA between the *Eco*RI and *Bam*HI restriction sites which surround the two α -globin genes.

Digestion of normal DNA with *Bgl*II or a combination of *Hpa*I and *Eco*RI generates two α -specific fragments, as these enzymes cut outside and between the two α genes (Fig. 1A, B) (refs 2, 6 and E. Whitelaw, personal communication).

When E.J. DNA was digested in this way the two normal bands were present and of equal intensity, but in addition, in each case a third less intense band, 3.7 kilobases long, was produced. This fragment must have been derived from a third α

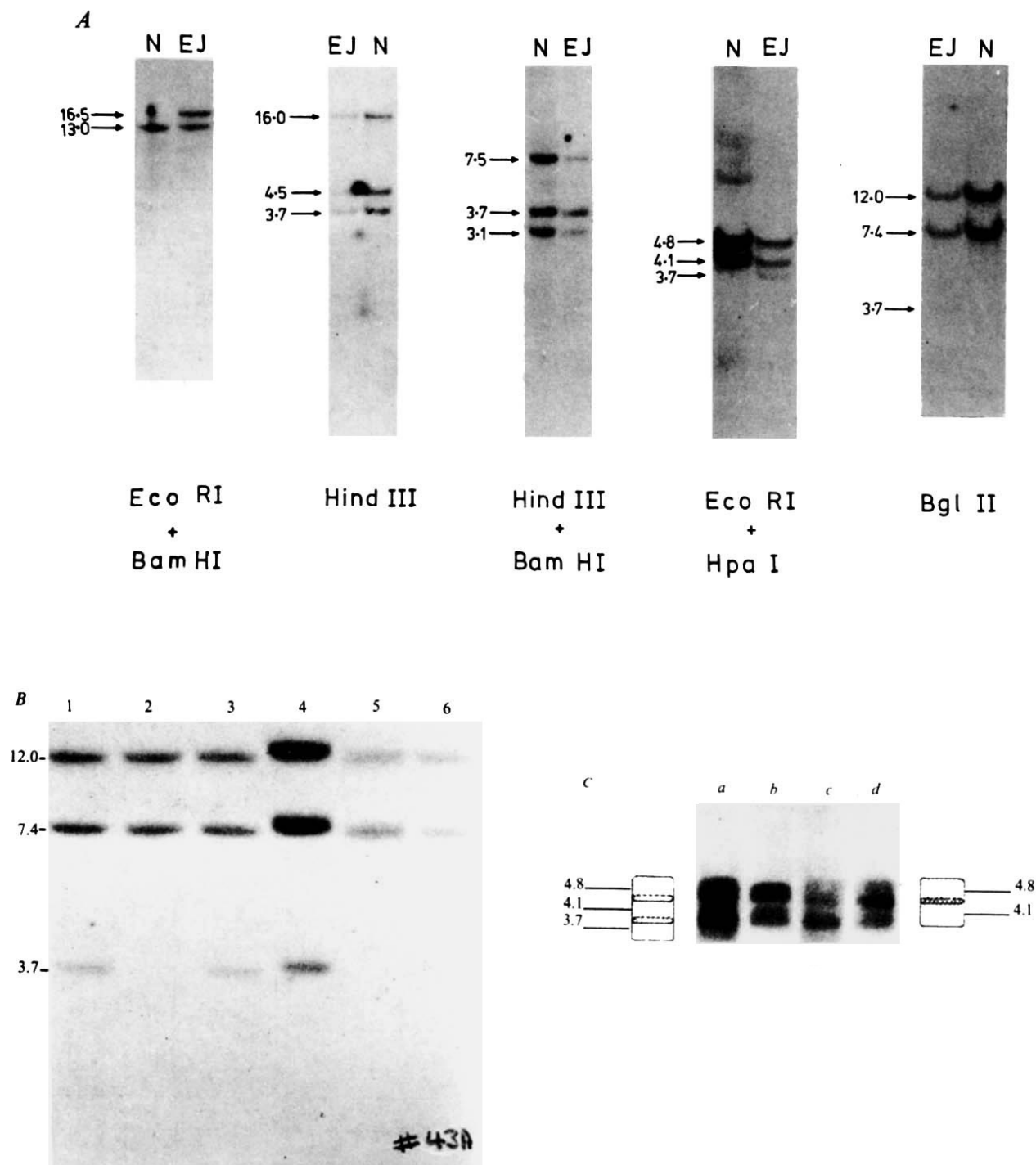


Fig. 1 A, Autoradiographs of DNA obtained from E.J. digested with various restriction enzymes. The sizes of DNA fragments are represented in kilobases. DNA was obtained by phenol chloroform extraction of peripheral blood or bone marrow⁹. Following restriction of 20 µg of DNA with 0.5–2.0 units of enzyme per µg DNA the samples were electrophoresed in 0.8% agarose and then transferred to nitrocellulose filters¹⁰. The DNA on the nitrocellulose filters was then hybridised to an α plasmid JW101 (ref. 11) which had been nick translated by the method of Maniatis *et al.* using ³²P-dCTP and ³²P-dATP (Radiochemical Centre) as radioactive labels¹². The filters were then washed in stringent conditions² and autoradiographed using Fuji-RX-X-ray film according to the method of Lasky and Mills¹³. An additional band (~10 kilobases) is present in the autoradiograph of the *Eco*RI/*Hpa*I double digest of normal DNA, representing an incomplete digest of the DNA which, in this case, was present in a greater amount than usual for such a digest. A faint (3.7 kilobase) band is present in the autoradiograph of the *Bgl*II digest of E.J. DNA in A. However, this band is clearly shown in B, lane 4. B, Autoradiographs of DNA obtained from D.J. (1), R.J. (2), C.J. (3), E.J. (4), normal (5), normal (6), which had been digested with *Bgl*II. The same pattern, consistent with the genotype *aaa/aa*, was present in DNA from E.J., D.J. and C.J. The intensity of the normal fragments compared with the new 3.7-kilobase fragment is approximately 2:1, which is the ratio predicted for this gene arrangement. C, Autoradiograph of fragments produced by *Hpa*I digestion of DNA isolated from the abnormal 16.5-kilobase fragment generated by cleavage of E.J. DNA with a combination of *Eco*RI and *Bam*HI as described in the text. The 13-kilobase fragment (channels b, d) produced two α -gene fragments 4.8 and 4.1 kilobases long, whereas the 16.5-kilobase fragment produced three α -specific fragments of 4.8, 4.1 and 3.7 kilobases (channels a, c). The bands produced in these autoradiographs are broader than normal because when the samples, which were cut out from the low-melting temperature agarose, had been restricted and then re-applied to agarose gel, the DNA was in a larger volume than is conventionally applied to agarose gel for electrophoresis.

Table 1 Phenotype and genotype of R.J., D.J. and C.J.

	RBC ($\times 10^{12} \text{ L}^{-1}$)	Hb (g dl^{-1})	Hct	MCV (fl)	MCH (pg)	Fe (mmol L^{-1})	TIBC (mmol L^{-1})	%HbA ₂	%HbF	α/β (counts)	α/β (S.A.)	α/β mRNA	Genotype
R.J.	4.8	14.2	0.399	82	29.5	21.5	88.5	2.50	0.67	1.01}	1.03}	2.5	$\alpha\alpha/\alpha\alpha$
D.J.	4.45	14.0	0.399	89	31.2	31.0	114.0	2.48	0.20	1.18}	1.06}	—	$\alpha\alpha\alpha/\alpha\alpha$
C.J.	4.70	14.3	0.409	86	30.4	19.0	79.5	2.47	0.29	1.11}	1.15}	2.9	$\alpha\alpha\alpha/\alpha\alpha$
										1.06}	1.24}		
										1.17}	1.00}		
										1.16}	0.99}		

Haematological and haemoglobin analysis was carried out using standard techniques⁷. All haematological studies were carried out on two separate occasions, with good agreement between duplicate results. E.J. had an unrelated haematological malignancy and therefore all studies were carried out on his haematologically normal sons R.J., D.J. and C.J. RBC, Red blood cell count; Hb, haemoglobin; Hct, haematocrit; MCV, mean corpuscular volume; TIBC, total iron-binding capacity; SA, specific activity.

gene present in the inserted DNA of the abnormal chromosome, with *Hpa*I and *Bgl*II sites 3.7 kilobases from their respective sites in the normal sequence between the two genes (Fig. 2). *Hind*III cuts normal DNA to the right and left and within the α -globin genes^{2,5,6} to produce three fragments of 16, 4.5 and 3.7 kilobases. The 16- and 4.5-kilobase fragments are cleaved by *Bam*HI to yield 7.5- and 3.1-kilobase fragments, respectively. Digestion of E.J. DNA with *Hind*III generated only the three normal fragments (Figs 1A, 2). The 16- and 4.5-kilobase *Hind*III fragments were both cleaved by *Bam*HI to yield normal-sized fragments. The *Bam*HI sites on either side of the two α -globin genes are thus positioned normally, and a new *Hind*III site exists inside the inserted DNA sequence, giving rise to a new fragment 3.7 or 4.5 kilobases long and therefore indistinguishable from the normal *Hind*III fragments. The positions of the *Hpa*I, *Bgl*II and *Hind*III sites indicate that the inserted DNA is a repeat of intergenic DNA and an α -globin gene. A similar conclusion was drawn from a double digest of E.J. DNA with *Hpa*I and *Hind*III; four fragments were observed, all corresponding to those found in normal DNA⁶.

In Fig. 2, the α -gene fragments generated by various restriction enzymes from a chromosome with three α -globin genes are compared with those from a normal chromosome. Although all the abnormal fragments can be generated from an abnormal chromosome without a third α -globin gene locus in the inserted

DNA, the intensity of the autoradiograph bands in the *Bgl*II digest and the *Eco*RI/*Hpa*I digest suggested that the abnormal chromosome does contain a third α -globin gene. In both digests, band intensities agree with the normal to abnormal band ratio of 2:1 predicted in Fig. 2.

Finally, E.J. DNA was digested with *Eco*RI plus *Bam*HI and the products fractionated in a low-melting temperature agarose gel. The large 16.5-kilobase fragment was separated from the normal 13.0-kilobase fragment by slicing the gel between and outside the calculated band positions. Both DNA fragments were redigested with *Hpa*I and fractionated in a normal agarose gel. The 13.0-kilobase fragment produced two α -gene fragments of 4.8 and 4.1 kilobases as expected. However, the

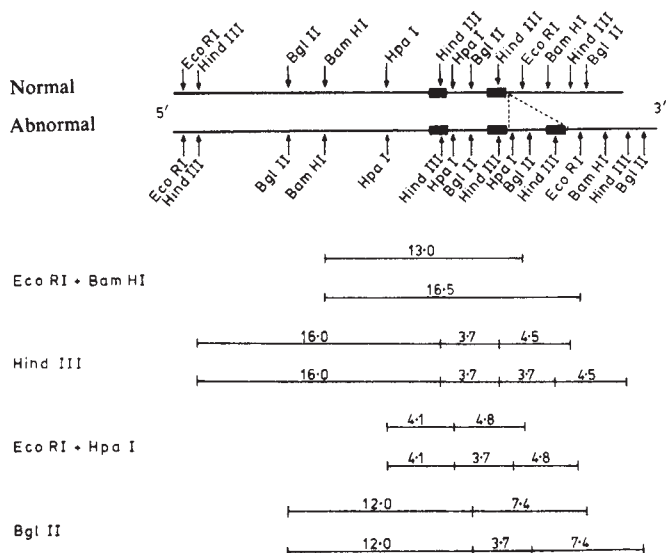


Fig. 2 The gene arrangement in E.J., C.J. and D.J. with two α genes on one chromosome and three on the other. The fragment sizes obtained from each chromosome following digestion with various restriction enzymes are shown below. The dotted lines indicate that the new arrangement contains an exact repeat of intergenic DNA and an α -globin gene. An identical map would be obtained if the inserted gene were placed to the right or left or between the normal duplicated α genes. The critical point is not the exact location of the inserted DNA but that it repeats precisely a part of the normal duplicated α -gene arrangement. As two fragments of equal size and one new fragment 3.7 kilobases long are produced by the enzyme *Bgl*II or a combination of *Hpa*I and *Eco*RI, the relative band intensity of 2:1 would be predicted from this model.

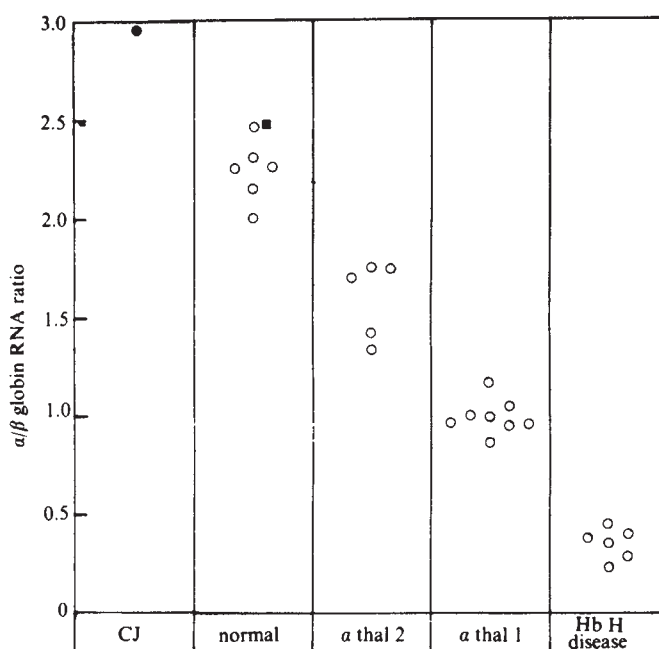


Fig. 3 α/β mRNA ratios in R.J. (■) and C.J. (●) compared with those of individuals with 4 (normal), 3 (α -thalassaemia 2; α thal 2), 2 (α -thalassaemia 1; α thal 1) and 1 (haemoglobin H disease; Hb H disease) functional α -globin genes. α/β mRNA ratios were determined as previously described¹. The range of α/β mRNA ratios in the cells of normal individuals is higher than those previously reported and the significance of this is discussed elsewhere¹. However, the importance of these results does not lie in the absolute values obtained for α/β -specific sequences but rather the relative values in the various α -thalassaemia syndromes. The α/β mRNA ratio in C.J. with the arrangement $\alpha\alpha\alpha/\alpha\alpha$ was about 20% higher than normal and falls within the range that would be predicted by extrapolation from the other groups if there was no compensation in the transcription of mRNA in individuals with five α -globin genes.

16.5-kilobase fragment produced three α -gene fragments of 4.8, 4.1 and 3.7 kilobases (Fig. 1C), proving conclusively that the abnormal chromosome contains three α -globin gene sequences instead of two. Such a situation would arise if the abnormal chromosome had arisen by an unequal crossing over

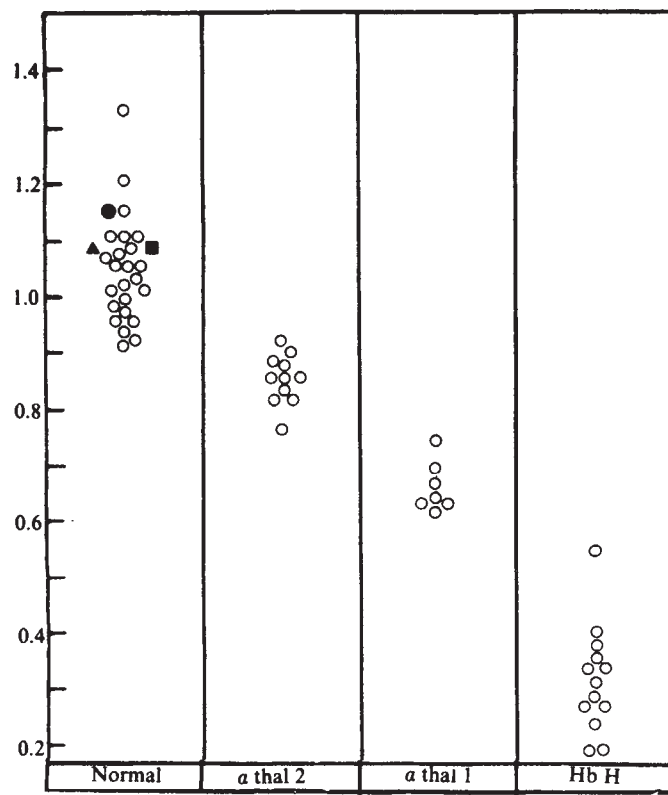


Fig. 4 α/β globin synthesis ratios in C.J. (●), D.J. (▲) and R.J. (■) compared with those individuals with varying numbers of functional α genes; α/β globin chain synthesis ratios were obtained as previously described¹⁴. The ratios obtained in 24 haematologically normal individuals, 11 subjects with α -thalassaemia 2, 7 subjects with α -thalassaemia 1, and 13 subjects with Hb H disease are shown. The ratios obtained in C.J., D.J. and R.J. are the average of two ratios obtained on separate occasions (Table 1) and fall within the range found in normal individuals.

between the α -globin genes, as proposed for the origin of the α -thalassaemia 2 genotype in several racial groups^{2,3}.

Haematological studies on D.J., C.J. and R.J. showed them to be normal (Table 1). Recently, α/β mRNA ratios in individuals with 4, 3, 2 and 1 functional α genes have been shown to fall into four non-overlapping groups¹. In R.J., who had the globin gene arrangement, $\alpha\alpha/\alpha\alpha$, the α/β mRNA ratio fell within the normal range (Fig. 3). However, in C.J., who had the $\alpha\alpha\alpha/\alpha\alpha$ globin gene arrangement, the α/β mRNA ratio was 20% higher than normal. Thus, it seems that when five α -globin genes are present, each is fully expressed at the mRNA level and there is no compensatory reduction in α mRNA output. However, the average α/β globin chain synthesis ratios in C.J. fell within the normal range (Table 1, Fig. 4).

Normal red cell precursors produce a slight excess of α chains, which are removed by proteolysis to achieve balanced α/β globin synthesis⁷. The fact that individuals with five α genes are haematologically normal and have no overall globin chain imbalance indicates that their red cell precursors can effectively compensate for an even greater excess of α -globin chain synthesis. However, the additional α gene might not always be functionally unimportant. Some cases of β -thalassaemia are unusually mild and there is genetic evidence that they are, in fact, compound heterozygotes for β -thalassaemia and 'silent β -thalassaemia', so called because it is not detectable in simple heterozygotes⁸. It is possible that some of these silent β -thalassaemias are, in fact, examples of the $\alpha\alpha\alpha/\alpha\alpha$ or $\alpha\alpha\alpha/\alpha\alpha\alpha$ genotype. Furthermore, one would predict that the basic thalassaemic nature of this genotype will only be revealed through its interactions with β -thalassaemia or β -chain haemoglobinopathies, such as Hb S.

This work was supported by the MRC. D.R.H. holds an MRC Training Fellowship. We thank Mrs H. Ayyub and Mrs B.

Aldridge for technical assistance, Dr B. Forget for plasmid JW101, and Dr C. Barton for bringing this family to our attention.

Note added in proof: Since this paper was submitted similar findings have been reported in individuals of Mediterranean background¹⁵.

Received 10 December 1979; accepted 25 February 1980.

1. Hunt, D. M. *et al.* *Br. J. Haemat.* (in the press).
2. Orkin, S. H. *et al.* *Cell* **17**, 33–42 (1979).
3. Phillips, J. A. III *et al.* *Blood* **54**, 1439–1445 (1979).
4. Baglioni, C. *Proc. natn. Acad. Sci. U.S.A.* **48**, 1880–1886 (1962).
5. Orkin, S. H. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5950–5954 (1978).
6. Embury, S. H. *et al.* *J. clin. Invest.* **63**, 1307–1310 (1979).
7. Weatherall, D. J. & Clegg, J. B. *The Thalassaemia Syndromes* 2nd edn (Blackwell, Oxford, 1972).
8. Schwartz, E. *New Engl. J. Med.* **281**, 1327–1333 (1969).
9. Gross-Bellard, M., Oudet, P. & Chambon, P. *Eur. J. Biochem.* **36**, 21–38 (1973).
10. Southern, E. M. *J. molec. Biol.* **98**, 503–517 (1975).
11. Wilson, J. T. *et al.* *Nucleic Acids Res.* **5**, 563–581 (1978).
12. Maniatis, T., Jeffrey, A. & Kleid, D. G. *Proc. natn. Acad. Sci. U.S.A.* **72**, 1184–1188 (1975).
13. Laskey, R. A. & Mills, A. D. *FEBS Lett.* **82**, 314–316 (1977).
14. Weatherall, D. J., Clegg, J. B. & Naughton, M. A. *Nature* **208**, 1061–1065 (1965).
15. Goossens, M. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **77**, 518–521 (1980).

Coordinately expressed members of two chorion multi-gene families are clustered, alternating and divergently orientated

C. Weldon Jones & Fotis C. Kafatos

Cellular and Developmental Biology, Biological Laboratories, Harvard University, Cambridge, Massachusetts, 02138

The chorion (eggshell) of silkmoths consists of more than 100 distinct proteins. Individual components are synthesised in characteristic and overlapping sequence, at early, middle, late or very late developmental periods of choriogenesis (which lasts 2 days in total)^{1,2}. Protein sequencing^{3,4} indicates that components which belong to two size classes, A and B, are encoded by two respective families of evolutionarily homologous genes (multi-gene families). Genetic analysis in *Bombyx mori*^{5,6} has shown that chorion genes are tightly linked in three clusters on a single chromosome ($n = 28$). Total chorion mRNA of the wild silkworm, *Antheraea polyphemus*, has been copied into double-stranded cDNA and cloned by recombinant DNA procedures⁷. Many distinct cDNA clones have been selected and identified as members of the A and B multi-gene families by detailed cross-hybridisation⁷ and sequencing⁸ analysis. Stringent conditions which minimise cross-hybridisation permit use of these clones as specific probes. In many cases, we know the protein encoded and the developmental period when the clone sequence can be found in cytoplasmic RNA^{7,8}. Thus, we may now use these cDNA clones as probes to determine how genes of known developmental properties and evolutionary history are arranged within the chorion chromosomal region. We report here that two cloned, 14-kilobase chromosomal DNA segments contain multiple chorion genes. In each segment, the genes represent both A and B families, but only one developmental class: one segment contains only middle-period A and B genes and the other only late-period A and B genes. In both segments, members of the two families are arranged in alternating sequence and divergent orientation.

A library of *A. polyphemus* chromosomal DNA clones was established⁹, using the charon 4 derivative of phage λ as vector¹⁰. Screening of approximately 180,000 phage plaques with total chorion cDNA permitted selection of a 'bookshelf' of 175 independent chromosomal clones containing chorion genes. After plaque purification, these clones were spotted on lawns of

Escherichia coli strain CSH18 and re-screened¹¹, using as probes nick-translated, characterised cDNA clones. Chromosomal clones were selected which hybridised strongly with two non-cross-reacting cDNA clones, even in stringent conditions (80 °C and 0.45 M NaCl, or approximately 10 °C below T_m for that probe length). Clone APc110 hybridised with two late sequences: the A-family sequence, pc18, and the B-family sequence, pc401. By contrast, clone APc173 hybridised with two middle-period sequences: the A-family sequence, pc292, and the B-family sequence, pc10. This suggested that chorion genes are clustered according to the developmental period when they are expressed, rather than according to the multi-gene family to which they belong. Preliminary screening of the library with additional probes supported that suggestion (data not shown).

For detailed analysis of gene arrangements, restriction maps of APc110 and APc173 were constructed by standard procedures (Fig. 1). Blot-hybridisation¹² experiments revealed in

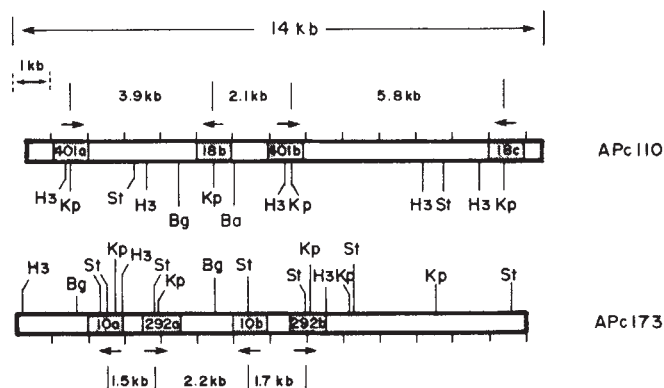


Fig. 1 Restriction maps of two chromosomal DNA clones from *A. polyphemus* carrying multiple chorion genes. Shading indicates the approximate location of genes numbered with reference to corresponding cDNA clones^{7,8}. Arrows show the direction of transcription. Clone APc110 contains only late-period genes, from two multigene families (pc18 = A; pc401 = B) and in opposite orientation. Clone APc173 contains only middle-period genes, again from two multi-gene families (pc292 = A; pc10 = B) and in opposite orientation. Locations of the genes were suggested by characteristic restriction sites known from the cDNA clones: *KpnI* for pc18, *SstI* for pc10, *KpnI* and *HindIII* for pc401, and *KpnI* and *SstI* for pc292. Confirmation came from detailed blot-hybridisation analysis using double enzyme digests. For example, the largest *KpnI* fragment of APc110 (~5.8 kilobases (kb)) contains both pc401 and pc18 sequences; after secondary digestion with *HindIII*, the left-most derivative (~3.6 kilobases) contains only pc401 sequences, the right-most derivative (~0.7 kilobases) contains only pc18 sequences, and the middle derivative (~1.5 kilobases) contains neither. Further confirmation came from blot-hybridisation experiments with 5'-specific or 3'-specific probes, which also defined the directions of transcription (see Fig. 2). One or more additional unmapped *SstI* sites exist between the two right-most *SstI* sites shown for APc173. Ba, *BamHI*; Bg, *BglII*; H3, *HindIII*; Kp, *KpnI*; St, *SstI*.

each case two non-contiguous copies of each of the represented cDNA sequences, in an alternating arrangement (BABA). Locations of the genes on the map were initially suggested by characteristic restriction sites known to be present in the clone cDNA sequences. These locations were verified by blot-hybridisation analysis of double enzyme digests (see Fig. 1 legend).

All four cDNA clones represented in these chromosomal segments have been sequenced. Thus, we could take advantage of centrally located restriction sites (*KpnI* for pc401 and pc18; *SstI* for pc10 and pc292) to generate probes representing either the 5' half or the 3' half of the corresponding mRNA. Blot-hybridisation of the chromosomal clones with these half-probes

defined the direction of transcription for each chromosomal gene (see, for example, Fig. 2). All experiments with half-probes gave a consistent directional map: A and B genes are transcribed in opposite orientations (arrows in Fig. 1), that is, A genes are transcribed from one DNA strand and B genes from the other.

The analysis with half-probes also established firmly that the blocks of each cDNA sequence represent two different gene copies, rather than one copy split in half by a very long intervening sequence. For example, in APc110, the pc401 half-sequences from left to right showed a 5'-3'-5'-3' progression, which is inconsistent with a single gene copy.

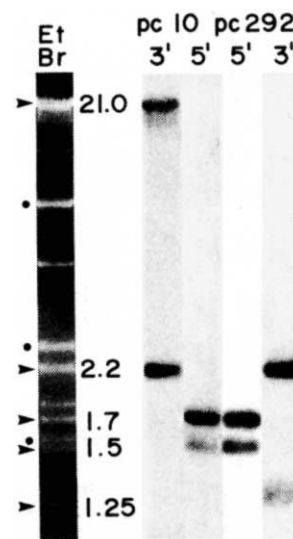


Fig. 2 Determination of transcriptional direction. The sequenced cDNA clones (ref. 8, unpublished observations, and S. G. Tsitilou, personal communication) were nick-translated and cleaved with *SstI* to generate separate probes specific for the 5' region or 3' region of each mRNA. Here blot-hybridisations of an *SstI* digest of APc173 are shown using such half-probes. The first track displays the ethidium bromide-stained total DNA pattern; arrowheads indicate DNA fragments which include chorion gene sequences (see the next four tracks), small dots indicate fragments which include moth DNA but no chorion sequences, and the unmarked fragments are derived solely from the charon 4 vector. The blot-hybridisation patterns show that the 5' regions of both pc10 and pc292 are found in the same two fragments (1.5 and 1.7 kilobases), whereas both 3' regions are only found together in a third fragment (2.2 kilobases); the 3' regions of pc10 and pc292 are also found individually, in two fragments of approximately 21.0 and 1.25 kilobases, respectively. These results are only consistent with divergent orientation of the genes (arrows in Fig. 1).

Blot-hybridisations of the chromosomal clones with total chorion cDNA revealed no additional chorion genes other than those shown in Fig. 1. However, sequencing of the left-hand end of APc110 revealed the presence of a very short sequence corresponding to the 5'-untranslated region of another pc18 gene (data not shown). Thus, two divergent AB pairs are found in each chromosomal clone (18a, 401a; 18b, 401b; 10a, 292a; 10b, 292b). We suspect that the right-hand copy of pc18 (18c in Fig. 1) is indicative of a third pair in APc110. The pairs are not part of invariant repeats; for example, the two *KpnI* fragments which contain the 3' ends of genes are unequal in length (3.9 and 5.8 kilobases), and also differ in several restriction sites.

The genes represented in APc110 correspond to two of the most abundant chorion mRNA sequences¹³. Do additional copies of these genes exist in the genome, and if so, are all copies found as divergent AB pairs? Figure 3 provides affirmative answers to both questions. Total moth DNA was restricted with *KpnI* and blot-hybridised with 5'-specific or 3'-specific probes

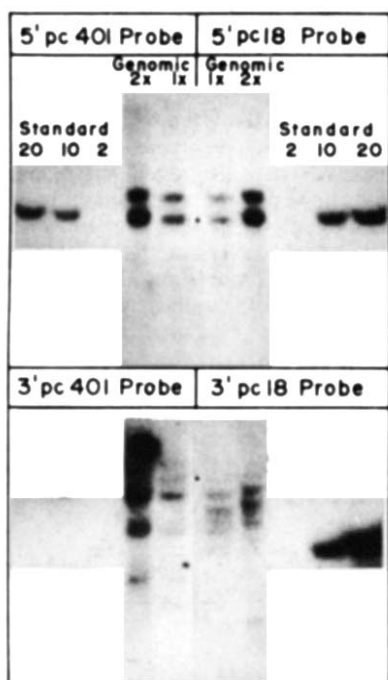


Fig. 3 Detection and organisation of repeated chorion genes in the moth genome. Total DNA from pooled developing pupae was digested with *KpnI*, electrophoresed, transferred to nitrocellulose¹² and hybridised with 5'- or 3'-specific ³²P-probes consisting of *KpnI* fragments from nick-translated cDNA clones pc401 and pc18; each probe contained the appropriate half of the chorion cDNA insert plus contiguous plasmid sequence. Hybridisation was at 75 °C in 0.45 M NaCl. Two aliquots of moth DNA were used in each experiment (1×, 10 µg; 2×, 20 µg). Reconstruction standards were included to estimate the gene multiplicity. They consisted of increasing amounts of a *HindIII/KpnI* digest of a subcloned *HindIII* fragment from APc110 (containing the entire *18b* gene and the 5' half of the *401b* gene; see Fig. 1), plus 10 µg sheared calf thymus DNA; only the hybridising *HindIII* to *KpnI* moth DNA fragments are shown. Each standard corresponds to the indicated number of gene copies per haploid genome (C = 1 pg), relative to the 1× moth DNA sample. Two major genomic bands hybridise with both 5' probes (top). When a mixture of the two probes is used, again only two bands are seen, confirming that the 5' ends of both types of genes hybridise to the same fragments (data not shown). The lower band (2.1 kilobases; dot) is expected from the map of APc110 (*KpnI* fragment containing the 5' ends of *18b* and *401b*; Fig. 1). The upper band (2.5 kilobases) indicates the existence of a second type of linked and divergent *401/18* pair in *A. polyphemus*. The combined intensities of the two bands, in comparison with the standards, indicate the existence of approximately 15 ± 5 *401/18* gene pairs per haploid genome (the gene *18* estimate is more dependable in this respect, as probe, standard and genomic fragment are co-terminous, whereas the standard lacks 130 of the 300 base pairs of *401* sequence represented in both probe and genomic fragment). At very long exposures, faint hybridisation to additional high molecular weight bands can be seen with both probes, possibly indicating limited polymorphism, single copies of additional *401/18* pairs, or cross-hybridisation with different but similarly organised genes of the same multi-gene families^{7,8}. Hybridisations with 3'-specific probes (bottom) reveal multiple bands, only some of which are identical for the two genes. This indicates that heterogeneity beyond the 3' ends of the genes is much greater than between the 5' ends. The expected 3.9-kilobase fragment (between *401a* and *18b*; Fig. 1) is indicated by a dot; the expected 5.8-kilobase fragment (between *401b* and *18c*) is seen in long exposures, but is smeared because of localised overloading of the gel (it co-migrates with a large number of unrelated *KpnI* genomic DNA fragments).

from cDNA clones pc401 or pc18, together with reconstruction standards. From the intensities of the blots, we estimate that approximately 15 ± 5 copies each of the *18* and *401* genes exist per haploid genome. With the 5' probes, only two major fragments hybridise, and they are the same for both *401* and *18*. The more intense fragment (which is found in approximately 10 copies per genome) corresponds to what is expected from clone APc110, whereas the other is longer by 0.4 kilobases. As these fragments were generated with *KpnI*, which cleaves within both *401* and *18* genes, and as each fragment contains the 5' ends of both *401* and *18*, we conclude that the multiple copies of these two genes are present in the genome as divergent AB pairs.

The map of APc110 shows length and restriction site variability 3' to the genes. As might be predicted, multiple bands of genomic DNA hybridise with each 3' probe, and some of these are different for the two probes. This confirms that greater variability is found in the long DNA segments containing the 3' ends of the genes than in the short DNA segments containing the divergent 5' ends. We conclude that APc110, and presumably also APc173, are representative: the multiple copies of four major chorion genes are arranged in pairs, each pair containing, in divergent orientation, coordinately expressed members of two different evolutionary families.

The co-expressed H3-H4 and H2A-H2B histone genes in *Drosophila* are arranged as divergent pairs within a repeat unit¹⁴, whereas in sea urchins the repeat includes only tandemly orientated genes¹⁵. Genes for the 70,000 molecular weight heat-shock protein are clustered in two loci, *87A* and *87C*, and include both tandemly orientated and divergent copies^{16,17}. Three apparently co-expressed ovalbumin-like genes are clustered in a tandem orientation¹⁸. Mammalian β -like globin genes are found in tandemly orientated pairs, arranged in progression according to their temporal expression¹⁹⁻²¹. Analysis of human deletions has shown that the developmental switching-off of one globin gene pair is controlled by a sequence element many kilobases away in the 3' direction (but located 5' to the next gene pair)¹⁹. Thus, coordinately expressed gene clusters may not be uncommon for eukaryotes. The chorion clusters are reminiscent of the histone genes in that they involve more than one sequence family. Clustering of co-expressed genes apparently does not reflect polycistronic transcription, because in some cases the genes are divergent, and because even the tandemly orientated globin genes seem to be transcribed individually^{22,23}. Although alternative interpretations are possible, clustering may be related to a gross level of regulation, such as availability for transcription in a particular type of differentiated cell, or at a certain developmental stage or physiological state, as a result of localised unravelling of chromatin²⁴; a visible manifestation of such a process might be the heat-shock-induced puffing of the *87A* and *87C* loci²⁵, each of which contains more than one gene copy. By contrast, divergent orientation might be related to finer-level regulation, such as precise temporal or quantitative coordination. In prokaryotes, divergent gene clusters are controlled by centrally located *cis*-acting regulatory elements²⁶; in addition, bi-directionally acting regulatory genes have been demonstrated in both prokaryotes and eukaryotes, although the transcriptional orientation of the flanking genes subject to regulation is not always known (reviewed in refs 26-28). Divergent organisation of repeated genes may also have interesting evolutionary implications: it may be that the unit of expansion or contraction in these multi-gene families is the gene pair, rather than individual genes. Functional interpretation of the intriguing gene arrangement reported here must await investigation of the generality of its occurrence among other chorion genes. It would be particularly interesting to know how extensive a chromosomal region limited to a single developmental period is, and whether it contains copies of more than two different genes.

This work was supported by grants from the NSF and NIH to F.C.K. C.W.J. was supported by an NSF predoctoral fellowship. Clone pc292 was characterised by Dr J. C. Regier and sequenced by Dr S. G. Tsitilou. Initial screening of many chromosomal clones with total chorion cDNA was done by Dr

G. K. Sim. We thank Drs T. Maniatis and W. J. Gehring for helpful discussions.

Received 6 November 1979; accepted 13 February 1980.

- Kafatos, F. C. *et al.* in *Results and Problems in Cell Differentiation* Vol. 8 (ed. Beerman, W.) 45-145 (Springer, Berlin, 1977).
- Paul, M. & Kafatos, F. C. *Dev. Biol.* **42**, 141-159 (1975).
- Regier, J. C., Kafatos, F. C., Goodfellow, R. & Hood, L. *Proc. natn. Acad. Sci. U.S.A.* **75**, 390-394 (1978).
- Rodakis, G. C. thesis, Univ. Athens (1978).
- Goldsmith, M. & Basehoar, G. *Genetics* **90**, 291-310 (1978).
- Goldsmith, M. & Clermont-Rattner, E. *Genetics* **92**, 1173-1185 (1979).
- Sim, G. K. *et al.* *Cell* **18**, 1303-1316 (1979).
- Jones, C. W., Rosenthal, N., Rodakis, G. C. & Kafatos, F. C. *Cell* **18**, 1317-1332 (1979).
- Maniatis, T. *et al.* *Cell* **15**, 687-701 (1978).
- Blattner, F. R. *et al.* *Science* **196**, 161-169 (1977).
- Benton, W. D. & Davis, R. W. *Science* **196**, 180-182 (1977).
- Southern, E. M. *J. molec. Biol.* **98**, 503-517 (1975).
- Kafatos, F. C., Jones, C. W. & Elstratidis, A. *Nucleic Acids Res.* **7**, 1541-1552 (1979).
- Lifton, R. P., Goldberg, M. L., Karp, R. W. & Hogness, D. S. *Cold Spring Harb. Symp. quant. Biol.* **42**, 1047-1051 (1977).
- Kedes, L. H. *Cell* **8**, 321-331 (1976); *A. Rev. Biochem.* **48** (in the press).
- Mirault, M.-E., Goldschmidt-Clermont, M., Artavanis-Tsakonas, S. & Schedl, P. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5254-5258 (1979).
- Holmgren, R., Livak, K., Morimoto, R., Freund, R. & Meselson, M. *Cell* **18**, 1359-1370 (1979).
- Royal, A. *et al.* *Nature* **279**, 125-132 (1979).
- Fritsch, E. F., Lawn, R. M. & Maniatis, T. *Nature* **279**, 598-603 (1979).
- Bernards, R., Little, P. F. R., Anison, G., Williamson, R. & Flavell, R. A. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4827-4831 (1979).
- Lacy, E., Hardison, R. C., Quon, D. & Maniatis, T. *Cell* **18**, 1273-1283 (1979).
- Tilghman, S., Curtis, P., Tiemeier, D., Leder, P. & Weissmann, C. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1309-1313 (1978).
- Kinniburgh, A. & Ross, J. *Cell* **17**, 915-921 (1979).
- Yamamoto, K. R. & Alberts, B. M. *A. Rev. Biochem.* **45**, 721-746 (1976).
- Ashburner, M. & Bonner, J. J. *Cell* **17**, 241-254 (1979).
- Otsuka, A. & Abelson, J. *Nature* **276**, 689-694 (1978).
- Arst, H. N. & MacDonald, D. W. *Nature* **254**, 26-34 (1975).
- Lewis, E. B. *Nature* **276**, 565-570 (1978).

Recombination of human influenza A viruses in nature

William J. Bean Jr

Division of Virology, St Jude Children's Research Hospital, Memphis, Tennessee 38101

Nancy J. Cox & Alan P. Kendal

WHO Collaboration Center for Influenza, Virology Division, Center for Disease Control, Atlanta, Georgia 30333

In 1977, a unique event occurred in the epidemiology of influenza when a virus of the influenza A (H1N1) subtype, similar to a virus that had occurred in 1950, reappeared and caused worldwide epidemics but did not replace the prevailing influenza A (H3N2) subtype¹⁻⁶. Consequently, the two viruses co-circulated throughout the world^{1,7} and mixed infection of some individuals with both virus strains was detected^{8,9}, raising the possibility that recombination between the two strains might affect the future epidemiological behaviour of influenza. Serological analysis of virus isolates from influenza outbreaks during the winter of 1978-79, however, failed to detect any antigenic hybrids (H3N1 or H1N2). The investigation described here, was therefore, undertaken to detect recombinants among recent isolates of the H1N1 and H3N2 serotypes, involving genes coding for other than the surface proteins by RNA-RNA hybridisation. We report here the genetic characterisation of recombinants of both antigenic types.

Sequence homologies of the genome RNA segments of influenza virus isolates were compared using a competitive RNA-RNA reassociation assay and by direct RNA-RNA hybridisation¹⁰. For the competitive reassociation assay double-stranded RNA probes were prepared by annealing ¹²⁵I-labelled individual RNA segments to homologous unlabelled cRNA. To

determine if a particular virus isolate contains an RNA segment related to that of a labelled reference RNA segment, the double-stranded RNA was melted and allowed to reanneal in the presence of increasing amounts of RNA from the virus isolate being tested. If the unlabelled viral RNA contains an RNA sequence identical to that of the ¹²⁵I-labelled segment, the unlabelled RNA will then compete for annealing sites on the complementary RNA and prevent the reannealing of the labelled RNA. If, however, the unknown virus RNA does not contain an homologous segment or contains one that is only partly homologous, then the RNA either would not compete for annealing sites on the cRNA or it would compete less efficiently and the labelled RNA would reanneal completely or partly with the cRNA.

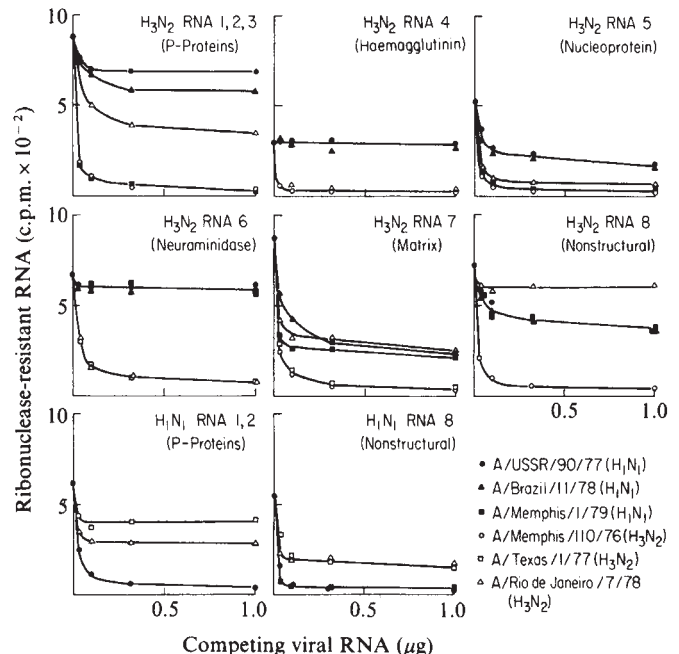


Fig. 1 Comparison of the RNA segments of recent influenza isolates with those of previous H3N2 and H1N1 strains by competitive hybridisation. The growth and purification of virus and the extraction of viral RNA has been described^{16,17}. Virion RNA (10 μg) from an H3N2 strain (A/Memphis/110/76) and an H1N1 strain (A/FW/1/50) was resolved into constituent genome segments by electrophoresis on a polyacrylamide-urea slab gel (22 × 14 × 0.15 cm) in 0.1 × Loening's buffer containing 7 M urea¹⁸, and 3% acrylamide cross-linked with 0.37% bis-acrylalcystamine¹⁹. This cross-linking reagent allows the gel to be solubilised by reduction. Electrophoresis was for 16 h at 25 °C with a constant voltage of 200 V. Following electrophoresis the gel was stained with ethidium bromide (5 μg ml⁻¹) and the RNA bands were visualised with a UV lamp and cut from the gel. The three largest RNA segments were not well resolved and were recovered in either one or two fractions. The gel slices were dissolved in 2 ml 0.05 M EDTA containing 10% 2-mercaptoethanol and 20 μg polyuridylic acid. (The poly-U acts as a carrier RNA during subsequent steps, but does not interfere with iodination¹⁷). The RNA and acrylamide were precipitated and washed with ethanol, resuspended in 2.5 ml sodium acetate buffer (pH 5.0) and the RNA was electrophoretically separated from the acrylamide in an Isco electrophoretic concentrator cell, operated at 50 V for 2 h. The RNA was recovered in 0.2 ml ethanol and precipitated. Each RNA segment was iodine labelled using 800 μCi ¹²⁵I as described by Comberford²⁰. The iodinated RNA was further purified by cellulose chromatography to remove contaminating ¹²⁵I-labelled material²¹. Complementary RNA was extracted from infected canine kidney cells as described by Etkind and Krug²². Double-stranded (ds) RNA probes were made by annealing each ¹²⁵I-labelled RNA segment to an excess of homologous complementary RNA and then removing remaining single-stranded RNA by digestion with nuclease S1 as described by Hay *et al.*²³. Each competitive reassociation assay used a constant amount of a ¹²⁵I-labelled ds probe and varying amounts of unlabelled RNA extracted from the strains being compared with the reference RNA segments. Reactions were carried out in 15 μl 2.3 × SSC, 50% formamide²⁴. The mixtures were heated at 90 °C for 90 s and then allowed to reanneal at 70 °C for 40 h. Acid precipitable radioactivity was determined by liquid scintillation counting following digestion with ribonuclease A and T₁ (ref. 24).

Table 1 Homology between RNAs of reference influenza virus and isolates, detected by quantitative RNA-RNA hybridisation

Complementary RNA from:	% Homology of cRNA with ³² P-labelled virion RNA segments					
	1	2	3	5	7	8
Reference virus		P		NP	M	NS
A/Texas/1/77(H3N2)	100	100	100	100	100	100
A/USSR/90/77(H1N1)	66	44	54	62	76	78
Recent isolates						
A/Shanghai/9/79(H3N2)	97	(105)		94	99	100
A/Singapore/333/79(H3N2)	110	(102)		98	101	97
A/Brazil/11/78(H1N1)	59	27	53	55	82	77
A/California/10/78(H1N1)	93	103	107	74*	84	73
A/Texas/8/78(H1N1)	92	97	98	79*	74	75
A/Texas/10/78(H1N1)	100	91	108	74*	78	80

Quantitative RNA-RNA hybridisations were carried out by annealing individual ³²P-labelled virion RNA segments from A/Texas/1/77 with homologous or heterologous complementary RNA as described by Scholtissek *et al.*¹⁰. ³²P-labelled virion RNA from the two reference influenza strains was extracted from virus grown in chick embryo kidney cells in the presence of ³²P and the individual genome RNA segments were isolated following polyacrylamide gel electrophoresis as described previously¹⁴. Complementary RNA was isolated from infected, cycloheximide-treated, chick embryo fibroblast cultures using the method of Stephenson *et al.*¹⁵. The labelled RNA segments were annealed to an excess of homologous or heterologous complementary RNA in 2×SSC and then heated at 75 °C in 1×SSC, 1% formaldehyde before digestion with ribonuclease¹⁰. Values are expressed as per cent of counts protected in the homologous reaction. Genes 4 and 6 were not analysed by hybridisation because they could be readily identified by antigenic analysis. Values greater than 100% reflect the error inherent in measurements of this type. The standard error of the procedure determined by repetitive tests of different RNA segments is ±5%.

* Value <100% probably due to cross-contamination of NP RNA (segment 5) with RNA of N2 neuraminidase (segment 6).

Initial RNA comparisons were made using probes prepared with RNA segments isolated from an H3N2 virus of the Victoria serotype (A/Memphis/110/76) and an H1N1 strain, A/FW/1/50, previously shown to be nearly identical to the 'USSR' virus³⁻⁵. Figure 1 shows results of competitive reassociation assays of various virus strains with the H3N2 probes and two representatives of the H1N1 probes. RNA extracted from the homologous strain or closely related strains gave essentially complete competition of the labelled probes. The heterologous RNAs competed to varying degrees with the different RNA segments. There was, for example, little cross-competition with the genes coding for the surface proteins of the prototype strains (genes 4 and 6), while genes 5 and 7 (nucleoprotein and matrix) showed much more cross-competition, indicating a greater conservation of the base sequence coding for these internal proteins, in agreement with the earlier findings of Scholtissek *et al.*⁴. In all cases, however, competition with RNA derived from the heterologous prototype strain was clearly distinguishable from the homologous competition.

One strain shown in Fig. 1, A/Memphis/1/79 (a recent H1N1 isolate), is clearly a recombinant, as four genes (1, 2, 3 and 5) coding for the three 90,000-molecular weight 'P' polypeptides and for the nucleoprotein are homologous with those of the H3N2 strain, while those coding for the surface proteins and genes 7 and 8, coding for the matrix and non-structural proteins, are homologous with those of the H1N1 strain. Comparable results (Table 1) were obtained by directly hybridising ³²P-labelled gene segments isolated from A/Texas/1/77 (H3N2) with complementary RNA isolated from cells infected with A/California/10/78 (H1N1) and other isolates from early outbreaks of influenza during the winter of 1978-79 in the US (A/Texas/8/78 and A/Texas/10/78). Three other H1N1 isolates obtained during the winter of 1978-79 (A/California/10/78; A/Berkeley/40/78; A/Georgia/64/78) were also tested by competitive hybridisation and shown to be identical recombinants (Table 2).

Many of the recent H1N1 strains mentioned above, which were shown to be recombinants between previous H1N1 and H3N2 viruses, are antigenically indistinguishable when tested with ferret sera or monoclonal antibodies (unpublished data) from A/Brazil/11/78, the prototype for a minor variant of the USSR strain predominant in South America during 1978 (ref.

11). Analysis of the RNAs of this prototype variant (Fig. 1, Table 1) indicated, however, that all of its gene segments were derived from an H1N1 strain.

To determine if the recent circulation of recombinant influenza viruses was limited to the above described H1N1 strains, the genetic composition of some recent H3N2 isolates¹² was also analysed. No evidence was found for the presence of genes of H1N1 origin in H3N2 strains isolated in south-east Asia in 1979 (Table 1). A South American H3N2 virus isolate (A/Rio de Janeiro/7/78), however, proved to have a different genetic composition. When RNA from this strain was tested in competitive reassociation assays (Fig. 1), only genes 4 and 6, coding for the surface proteins, of the H3N2 strain were shown to have homologous counterparts in this virus. The RNA did not contain sequences homologous to the other genes of either the H1N1 or H3N2 prototypes, although gene 5 (nucleoprotein) seemed to be more closely related to the corresponding H3N2 gene than to that of H1N1.

In an effort to determine the origin of the genes coding for the non-surface proteins of this virus, ¹²⁵I-labelled probes were prepared from its RNA segments and these were used in competitive reassociation assays with RNA from a series of 20 viruses. These included representatives of the major human influenza A viruses and selected strains from lower animals, as well as an influenza B virus. Mixtures of H3N2 and H1N1 RNA were also tested to rule out the possibility that this virus might actually be a mixture with two copies of each gene. With each segment only RNA from A/Rio de Janeiro/778 was effective in competition assays. None of the other strains tested were found to contain completely homologous RNA.

Our findings that the H1N1 virus prevalent in the US during the winter of 1978-79 contained four genes from an H3N2 virus, are similar to those of Young and Palese¹³, who studied these viruses by oligonucleotide and peptide mapping. Evidently, this recombinant had a distinct selective advantage over the parental strains. Possibly, the immune status of the population favoured the surface proteins of the H1N1 variant, while the replication complex of the H3N2 strain was better adapted for infection and transmission. The earliest examples of these recombinant viruses so far obtained were isolated in November and December of 1978 (Table 2). As these viruses

Table 2 Summary of gene derivation of recent H1N1 and H3N2 virus isolates analysed

	RNA segments							
	1	2	3	4	5	6	7	8
	P			HA	NP	NA	M	NS
H1N1 strains								
A/Brazil/11/78								
8 May 78	H1	H1	H1	H1	H1	H1	H1	H1
A/Berkeley/40/78								
29 Nov 78								
A/Calif/10/78								
10 Dec 78								
A/Texas/8/78								
28 Nov 78	H3	H3	H3	H1	H3	H1	H1	H1
A/Texas/10/78								
12 Dec 78								
A/Memphis/1/79								
2 Feb 79								
A/Ga/64/79								
19 Feb 79								
H3N2 strains								
A/Rio de Janeiro/7/78								
8 May 78	?	?	?	H3	?	H3	?	?
A/Singapore/333/79								
7 April 79	H3	H3	H3	H3	H3	H3	H3	H3
A/Shanghai/9/79								
16 April 79								

H1, RNA segment homologous with corresponding gene of A/USSR/90/77 (H1N1).

H3, RNA segment homologous with corresponding gene of recent H3N2 isolates of the Victoria or Texas serotypes.

?, RNA segment not closely related to H1N1 or H3N2 strains or any other influenza strain yet tested.

are antigenically identical to the H1N1 variant, which was isolated about 6 months earlier, our finding that A/Brazil/11/78 contained all the genes from the USSR strain indicates that the drift of the H1 haemagglutinin from A/USSR/80/77 to A/Brazil/11/78 has occurred independently of recombination.

The significance of the finding that the H3N2 virus A/Rio de Janeiro/7/78, which contains genes coding for surface proteins identical to those of other recent H3N2 isolates, but derives all other genes from an unknown source remains uncertain. This isolate was chosen for study primarily because it came from an outbreak in the same country and in the same week as A/Brazil/11/78, and therefore was of interest to us as a potential recombinant with H1N1 genes. This virus and others from the outbreak also have the unusual property of cross-reacting equally with both the Texas and Victoria serotypes, as recently found for a small number of other H3N2 isolates⁷. It was isolated in chicken embryos and egg passaged several times before use in this study. Neither the hybridisation results nor subsequent cloning have given any evidence that it is a mixture of viruses. The unknown genes of A/Rio de Janeiro/7/78 may have been derived from a previously unrecognised human influenza virus or from some animal influenza that was not tested

in the experiments described here. In either case, the identity of only the haemagglutinin and neuraminidase genes of A/Rio de Janeiro/7/78 with those of recent H3N2 strains suggests that this unusual virus may be a further example of a recombinant virus infecting man.

We thank Mr Raymond Wilson for technical assistance and Dr Peter Palese for prepublication information. This work was supported, in part, by grant AI 14293 from the National Institute of Allergy and Infectious Disease, by Cancer Center Support (CORE) grant CA 21765 from the National Cancer Institute and by ALSAC.

Note added in proof: The unusual H3N2 virus strain described above (A/Rio de Janeiro/78) has recently been analysed by J. Young and P. Palese and shown to be genetically identical to a laboratory-derived recombinant of a Victoria strain in the A/PR/8/34 strain (personal communication). This suggests that A/Rio de Janeiro/78 may not be a natural "recombinant". We cannot be certain, however, whether the isolation of this virus and several other identical isolates was due to actual infection with a laboratory-derived strain or to laboratory contamination of clinical samples. Epidemiological studies to resolve this question are underway.

Received 3 December 1979; accepted 13 February 1980.

1. WHO Wkly Epidemiol. Rec. No. 4, 25-28 (1979).
2. Zhdanov, V. M. et al. *Lancet* **i**, 294-295 (1978).
3. Kendal, A. P., Noble, G. R., Skehel, J. J. & Dowdle, W. R. *Virology* **89**, 632-636 (1978).
4. Scholtissek, C., von Hoyningen, V. & Rott, R. *Virology* **89**, 613-617 (1978).
5. Nakajima, K., Desselberger, U. & Palese, P. *Nature* **274**, 334-339 (1978).
6. US Dept of Health, Education and Welfare. *Morbidity and Mortality Weekly Rep.* **28**, 166-167 (1979).
7. Kendal, A. P. et al. *Am. J. Epidemiol.* **110**, 449-461 (1979).
8. Kendal, A. P. et al. *Am. J. Epidemiol.* **110**, 462-468 (1979).
9. Kendal, A. P. et al. *Abstr. A. Meet. Am. Soc. Microbiol.* 304 (1979).
10. Scholtissek, C., Harms, E., Rohde, W., Orlich, M. & Rott, R. *Virology* **74**, 332-344 (1976).
11. Webster, R. G., Kendal, A. P. & Gerhard, W. *Virology* **96**, 258-264 (1979).
12. US Dept of Health, Education and Welfare. *Morbidity and Mortality Weekly Rep.* **28**, 348 (1979).
13. Young, J. F. & Palese, P. *Proc. natn. Acad. Sci. U.S.A.* **76**, 6547-6551 (1979).
14. Kendal, A. P., Cox, N. J., Galphin, J. C. & Maassab, H. F. *J. gen. Virol.* **44**, 443-452 (1979).
15. Stephenson, J. R., Hay, A. J. & Skehel, J. J. *J. gen. Virol.* **36**, 237-248 (1977).
16. Hinshaw, V. S., Bean, W. J., Webster, R. G. & Easterday, B. C. *Virology* **85**, 51-62 (1978).
17. Bean, W. J., Sriram, G. & Webster, R. G. *Analyt. Biochem.* **102**, 228-232 (1979).
18. Floyd, R. W., Stone, M. P. & Joklik, W. K. *Analyt. Biochem.* **59**, 599-609 (1974).
19. Hansen, J. N. *Analyt. Biochem.* **76**, 37-44 (1976).
20. Commerford, S. L. *Biochemistry* **10**, 1943-1999 (1971).
21. Franklin, R. M. *Proc. natn. Acad. Sci. U.S.A.* **55**, 1504-1511 (1966).
22. Etkind, P. R. & Krug, R. M. *J. Virol.* **16**, 1464-1475 (1975).
23. Hay, A. J., Lomniczi, B., Bellamy, A. R. & Skehel, J. J. *Virology* **83**, 337-355 (1977).
24. Bean, W. J., Jr & Simpson, R. W. *J. Virol.* **18**, 365-369 (1976).

***E. coli* RNA polymerase promoters on superhelical SV40 DNA are highly selective targets for chemical modification**

Patricia Hale, Richard S. Woodward & Jacob Lebowitz

Department of Microbiology and Comprehensive Cancer Center, University of Alabama in Birmingham, Birmingham, Alabama 35294

Supercoiled DNAs are required or preferred in a number of biological processes such as repair and transcription. This requirement may be due to special properties which are present in supercoiled (FI) DNAs but not in the corresponding allomorphic forms. If sufficiently supercoiled, FI DNAs seem to contain single-stranded regions that are accessible to single-strand specific reagents and endonuclease¹⁻³. The free energy associated with superhelix formation may also be used to drive ligand-DNA or protein-DNA interactions³⁻⁷. *In vitro* studies have shown that superhelical DNA is a better template for transcription than its corresponding allomorphic forms⁸⁻¹⁵, and this is due probably to a change in the initial binding affinity of the RNA polymerase at promoter sites¹⁵⁻¹⁸. Previous work in our laboratory has suggested that the single-strand specific probe *N*-cyclohexyl-*N'*-β-(4-methylmorpholinium) ethyl carbodiimide (CMC), which almost totally inhibits transcription of SV40 and PM2 DNA^{14,18}, is acting preferentially at promoter sites¹⁸⁻²⁰. We report here experiments to determine the selectivity of CMC for promoters on superhelical SV40 DNA. After limited reaction with CMC there was substantially less *Escheri-*

chia coli RNA polymerase bound to reacted SV40 DNA than to the unmodified DNA. Marked inhibition of transcription from the modified DNA demonstrated that CMC is highly selective for *E. coli* RNA polymerase promoters in superhelical DNA, suggesting that the introduction of superhelical turns may alter the parameters by which the enzyme interacts with DNA through promotor recognition and opening of the DNA.

E. coli RNA polymerase was bound to supercoiled SV40 DNA at concentrations exceeding those for saturation as determined by previous experiments¹⁵. Complexes were cross-linked, purified and cleaved with *HpaII* (Fig. 1). *HpaII* endonuclease cleaves SV40 at 0.735 map units; this area does not contain any *E. coli* RNA polymerase binding sites¹⁵. Typical SV40 DNA complexes with *E. coli* RNA polymerase are shown in Fig. 1. This figure also demonstrates the efficient removal of excess polymerase and glutaraldehyde by Sepharose 6B column chromatography (full details will be published elsewhere). *E. coli* RNA polymerase binding sites were precisely mapped on over 400 DNA-RNA polymerase complexes in three separate experiments. Figure 2 gives map measurements of ~200 molecules in one experiment. *E. coli* RNA polymerase binding sites were consistently observed at positions 0.13, 0.18-0.21, 0.47, 0.52-0.53, 0.60-0.62, 0.65, 0.82, 0.925 and 0.975 map units on an SV40 map using the single *EcoRI* cleavage site as map position 0.0. As the number of bound polymerase molecules increased, it became more difficult to spread the DNA sufficiently for this type of map measurement. Most complexes containing nine molecules of RNA polymerase per DNA molecule were observed in a collapsed form due to polymerase-polymerase interactions. Therefore, most molecules used in map measurements contained five or six polymerase per DNA molecule. However, the localisation of binding sites on a sub-population of molecules that contained nine molecules per

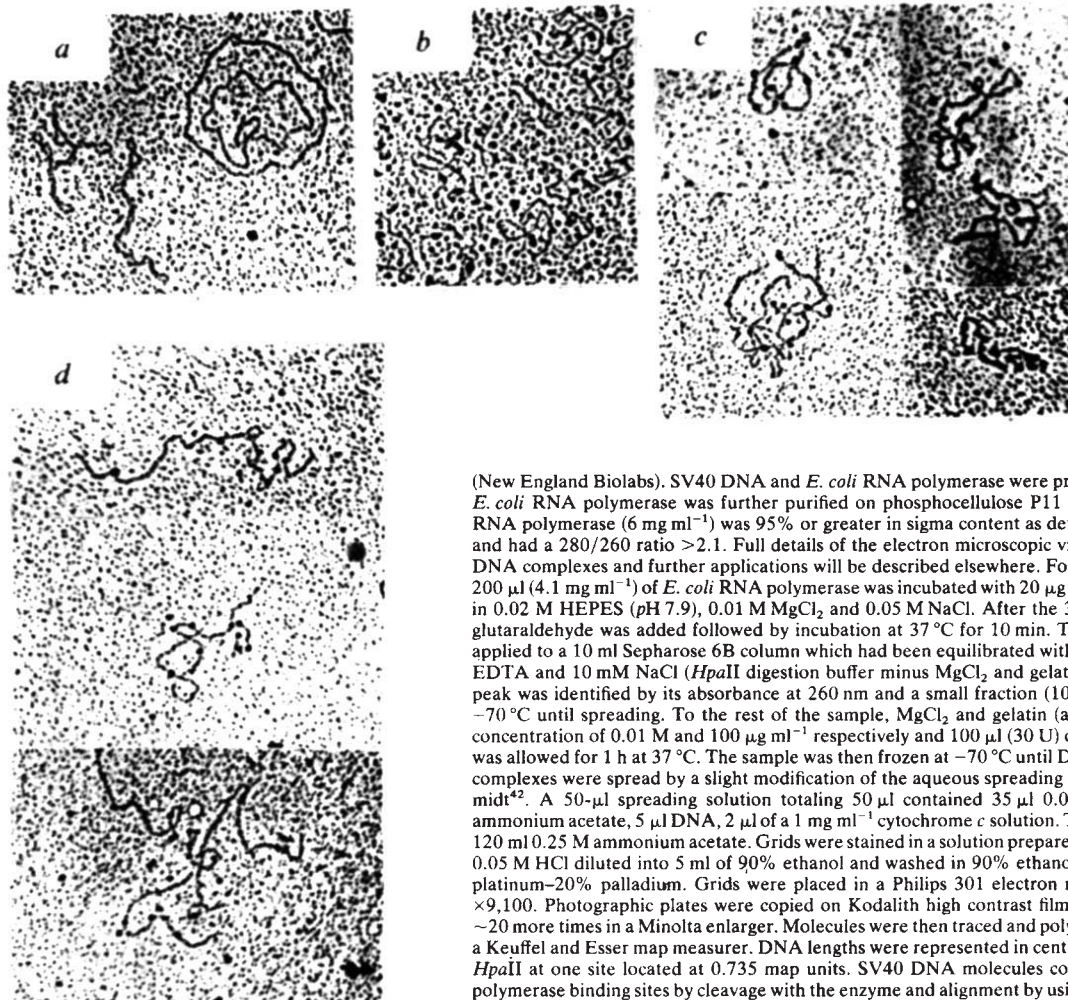


Fig. 1 Enhanced visualisation of glutaraldehyde-fixed RNA polymerase-SV40 DNA complexes by removal of excess RNA polymerase and fixative using Sepharose 6B column chromatography. Electron micrographs of: *a*, superhelical SV40 DNA FI; *b*, *E. coli* RNA polymerase-SV40 DNA complexes before Sepharose 6B chromatography; *c*, *E. coli* RNA polymerase-SV40 DNA complexes after Sepharose 6B chromatography; *d*, *E. coli* RNA polymerase-SV40 DNA complexes cleaved with *HpaII* restriction endonuclease

(New England Biolabs). SV40 DNA and *E. coli* RNA polymerase were prepared as described previously¹⁵. *E. coli* RNA polymerase was further purified on phosphocellulose P11 as described previously⁴⁰. *E. coli* RNA polymerase (6 mg ml⁻¹) was 95% or greater in sigma content as determined by rifampin inhibition⁴¹ and had a 280/260 ratio >2.1. Full details of the electron microscopic visualisation of RNA polymerase-DNA complexes and further applications will be described elsewhere. For binding of polymerase to DNA, 200 µl (4.1 mg ml⁻¹) of *E. coli* RNA polymerase was incubated with 20 µg of SV40 DNA at 37 °C for 30 min in 0.02 M HEPES (pH 7.9), 0.01 M MgCl₂ and 0.05 M NaCl. After the 30 min incubation, 6.25 µl of 8% glutaraldehyde was added followed by incubation at 37 °C for 10 min. The sample was then immediately applied to a 10 ml Sepharose 6B column which had been equilibrated with 10 mM HEPES (pH 7.4), 1 mM EDTA and 10 mM NaCl (*HpaII* digestion buffer minus MgCl₂ and gelatin). The DNA-RNA polymerase peak was identified by its absorbance at 260 nm and a small fraction (100 µl) was removed and frozen at -70 °C until spreading. To the rest of the sample, MgCl₂ and gelatin (autoclaved) were added to a final concentration of 0.01 M and 100 µg ml⁻¹ respectively and 100 µl (30 U) of *HpaII* was added and digestion was allowed for 1 h at 37 °C. The sample was then frozen at -70 °C until DNA and DNA-RNA polymerase complexes were spread by a slight modification of the aqueous spreading technique described by Kleinschmidt⁴². A 50-µl spreading solution totaling 50 µl contained 35 µl 0.001 M EDTA (pH 7.5), 5 µl 5 M ammonium acetate, 5 µl DNA, 2 µl of a 1 mg ml⁻¹ cytochrome *c* solution. The hypophase solution contained 120 ml 0.25 M ammonium acetate. Grids were stained in a solution prepared with 5 µl 0.05 M uranyl acetate, 0.05 M HCl diluted into 5 ml of 90% ethanol and washed in 90% ethanol, dried and shadowed with 80% platinum-20% palladium. Grids were placed in a Philips 301 electron microscope and photographed at ×9,100. Photographic plates were copied on Kodak high contrast film (Eastman-Kodak) and enlarged ~20 more times in a Minolta enlarger. Molecules were then traced and polymerase positions measured using a Keuffel and Esser map measurer. DNA lengths were represented in centimetres. SV40 DNA is cleaved by *HpaII* at one site located at 0.735 map units. SV40 DNA molecules could be orientated for location of polymerase binding sites by cleavage with the enzyme and alignment by using the information obtained from polymerase binding to *HinII*-III restriction fragment¹⁵. Regions from 0.05 to 0.01, 0.03 to 0.04, 0.65 to 0.75 and 0.93 to 0.965 map units did not contain polymerase binding sites¹⁵. This allowed only one possible orientation for essentially all of the molecules that were measured. The number of polymerase molecules bound to each DNA molecule was counted directly from the photographic plates. Before spreading, relaxed circular SV40 DNA was added to DNA-RNA polymerase complexes in a concentration of 0.5 µg ml⁻¹. This DNA acted as a marker for determining full-length DNA-RNA polymerase complexes for mapping of polymerase binding sites. At this concentration, at least 4 or 5 measurable circular molecules of SV40 DNA were found on each photographic plate.

DNA molecule and were sufficiently spread for map measurement gave a similar result (data not shown).

The total number of *E. coli* RNA polymerase molecules bound to each SV40 FI DNA molecule was determined for over 400 complexes at saturating concentrations of *E. coli* RNA polymerase (see Fig. 3a for data on molecules used for map measurements in Fig. 2). A maximum was observed at nine polymerase molecules per DNA molecule. The results indicate that in the supercoiled form of SV40 DNA there are nine specific RNA polymerase binding sites which may be occupied at saturating concentrations of the enzyme. No preferential binding to any particular site was observed, although increased stability from dissociation has been observed for the binding sites at 0.13 and 0.18–0.21 map units (P. Hale, unpublished observation).

The apparent equilibrium reactivity of the single-strand regions in the superhelical form is between 108 and 121 CMC molecules in the conditions used^{19,21}. No detectable reaction of nicked SV40 DNA was observed and saturating ethidium bromide density measurements did not detect any significant decrease in superhelical density²¹.

We have studied polymerase binding to, and template activity of, SV40 FI DNA which has been modified by 7 and 15 min of CMC reaction. The mean extent of carbodiimide binding after 15 min of reaction is 10.7 molecules CMC per molecule SV40 as determined by buoyant density analysis²² (Fig. 4). The linear nature of the kinetics allows interpolation to 7 min of reaction;

after fitting the data to a straight line, the mean reactivity at 7 min was found to be 5.0 CMC molecules per SV40 DNA molecule.

Evaluation of the extent of RNA polymerase binding by electron microscopy as well as the inhibition of transcription in

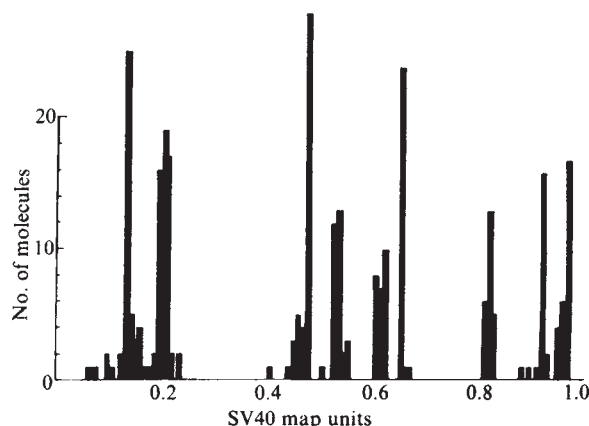


Fig. 2 Map location of *E. coli* RNA polymerase binding sites on native SV40 DNA FI as determined by electron microscopy. Samples were prepared as described in Fig. 1.

Table 1 Comparison of RNA polymerase binding and transcription between untreated SV40 FI DNA and 7 and 15 min CMC-reacted DNA

A	B	C	D	E	F	G
CMC reaction (min)	Moles CMC per mole DNA	Mean moles polymerase per mole DNA	Moles CMC per polymerase binding site	% Inhibition of polymerase binding	% Inhibition of transcription -Rif	% Inhibition of transcription +Rif
0	0	7.1	0	0	0	0
7	5.0	—	0.56	—	35%	54%
15	10.7	4.5	1.2	37%	61%	88%

A, SV40 DNA I (25 µg) was reacted with a 500 molar excess of CMC and purified from excess reagent as previously described¹⁸. B, The molecules of CMC bound per molecule of SV40 DNA were determined by buoyant density analysis as shown in Fig. 4. The 7-min reaction was interpolated from the data. C, The mean number of RNA polymerase molecules per SV40 DNA was determined from the data in Fig. 3. D, The number of CMC molecules per RNA polymerase site was calculated by dividing column B by nine, the total number of RNA polymerase binding sites on untreated SV40 DNA. E, The per cent inhibition of binding was calculated from the data of column C. F and G represent transcription assays of RNA polymerase-SV40 DNA complexes before preparation for electron microscopy. Transcription assays were performed as previously described^{15,18}, with DNA and polymerase incubated together for 10 min before the addition of nucleotide triphosphates (NTP). In transcription assays containing CMC-reacted DNA, DNA was incubated in transcription buffer 10 min before polymerase addition as previously described¹⁸. CMC reactions were repeated 4 times with per cent inhibition of transcription, with or without rifampin, showing reproducibility of $\pm 2\%$. Rifampin, 4 µg ml⁻¹, was added with NTP in column G.

the presence and absence of rifampin should reveal whether extremely limited CMC modification is specific for promoter sites. The electron microscopy technique described was used to study the effects of 15-min CMC modification of the DNA on RNA polymerase binding (Fig. 3b). Maximum binding of RNA polymerase was observed at five polymerase molecules per DNA molecule. This loss of RNA polymerase binding is compared with the inhibition of transcription in Table 1. It can be seen that an average of 1.2 CMC molecules per polymerase site is sufficient to produce a 37% loss of binding as judged by electron microscopy (Table 1, D, E). Transcription inhibition of 61 and 88% occurred in the absence and presence of rifampin respectively (Table 1, F, G). Although electron microscopic mapping was not carried out on the 7 min CMC-reacted DNA (5.0 CMC per SV40 DNA) the lower inhibition of transcription is consistent with the data obtained for the 15-min CMC-reacted DNA (10.7 CMC per SV40 DNA). These results show that the *E. coli* RNA polymerase promoters on SV40 DNA FI exist in a state, or states, readily accessible to extremely limited modification by a reagent highly selective for unpaired bases. Rifampin resistance measures the capacity to initiate rapidly a ternary complex for transcription. Previous work with native SV40 DNA has shown that there is essentially no difference in transcription with or without rifampin¹⁵. Here we observe that on very limited CMC modification rifampin causes a further decrease in transcription of about 20% (Table 1, F, G) compared with transcription without rifampin. Hence it seems that CMC modification enhances rifampin sensitivity by weakening promoters. This is supported by earlier data for 2-h (CMC-reacted) SV40 DNA tested for rifampin sensitivity. In these experiments, sensitivity to the drug was measured by differences in template saturation at different rifampin concentrations¹⁸. The saturation point was decreased by 26 and 47% for 2 and 4 µg ml⁻¹ rifampin, respectively. Unmodified SV40 DNA FI is insensitive to increase in rifampin from 2 to 8 µg ml⁻¹ (ref. 15). The capacity of CMC to weaken or block RNA polymerase binding could depend on the nature of individual promoters. This has been tested by mapping the binding sites for one of the modified SV40 DNA samples.

Figure 5 shows the results of mapping RNA polymerase molecules bound to 15 min CMC-reacted SV40 FI DNA. When these results are compared with those for the untreated DNA (Fig. 2), it can be seen that the location of the sites is very similar on both the CMC-reacted and unmodified DNA. Nine sites are observed at map positions 0.11, 0.18–0.19, 0.46, 0.54, 0.60, 0.63, 0.83, 0.925 and 0.98. The binding sites were observed to be more heterogeneous on the CMC-reacted DNA, this may be due to the decrease in rifampin-resistant tight binding sites on the CMC-modified DNA. This would be anticipated if CMC modification in or near a given promoter could occur at different locations leading to a variation in binding position around the promoter site. No specific population of sites was observed to be particularly sensitive to CMC reaction or preferentially lost by a 15-min reaction of the DNA with CMC.

Also note that CMC modification increases RNA polymerase binding around the 0.30 and 0.40 map positions to a small extent. Recent assays of unmodified SV40 DNA restriction fragments retained on nitrocellulose filters by RNA polymerase show that these regions do contain weak affinity for the enzyme (P.T. Chan, unpublished observation).

The use of a slow-reacting reagent^{19,21} (108–124 molecules CMC/SV40) and only short reaction times has allowed us to limit the reaction to approximately one CMC per promoter. This situation precludes propagation of denaturation into RNA polymerase binding regions by modification at neighbouring sites. No significant loss of superhelical turns can occur with only

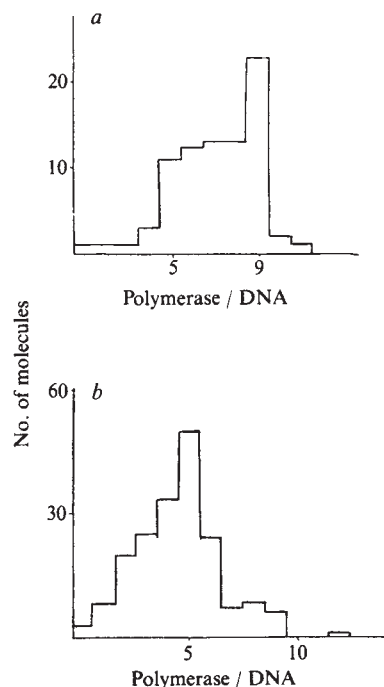


Fig. 3 The respective distributions of the number of *E. coli* RNA polymerase molecules per SV40 FI DNA for: a, unmodified; and b, 15 min. CMC-modified DNA. Supercoiled SV40 FI native or CMC-modified DNA was incubated with *E. coli* RNA polymerase, cross-linked with glutaraldehyde, chromatographed on Sepharose 6B and cleaved with *Hpa*II as described in Fig. 1. The resulting complexes were then spread for electron microscopy as described. The number of RNA polymerase molecules per DNA molecule were counted directly from the photographic plates. The reaction of SV40 FI DNA with CMC was performed at a 500-fold molar excess of CMC to nucleotides for 15 min at 37°C and excess reagent was removed by chromatography on Biogel A5m as described previously¹⁸. CMC-modified DNA was incubated with bovine serum albumin (0.5 mg ml⁻¹) for 10 min at 37°C before polymerase addition to absorb residual CMC¹⁸.

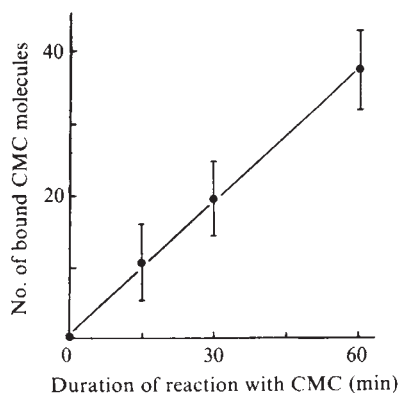


Fig. 4 Extent of CMC modification of SV40 DNA I as a function of time. The amount of CMC bound in the conditions described in Fig. 3 was evaluated using the equation of Bauer and Vinograd²²

$$\nu' = \frac{\theta(\bar{\nu}_3 + \Gamma') - (1 + \Gamma')}{(1 - \bar{\nu}_4\theta)}$$

where ν' is g of reagent bound per g of DNA; θ , the buoyant density of the modified DNA; $\bar{\nu}_3$, the partial specific volume of anhydrous DNA (0.479); Γ' , the net hydration of the modified DNA; and $\bar{\nu}_4$ the partial specific volume of CMC (0.814)¹. Γ' was evaluated by the equations developed by Bauer and Vinograd²². The moles of CMC per mol of nucleotide was obtained by multiplying ν' by 430/423, the respective molecular weights of a Cs nucleotide and CMC. The buoyant density of SV40 DNA was determined from a G-C content of 0.4067 from the full sequence data²⁸ using the relationship $\theta = 0.0988 (G-C) + 1.6541$. This gives a θ value for SV40 DNA of 1.6944 g cm⁻³. The above unpublished equation for θ against (G-C) is relative to a θ value for *E. coli* DNA of 1.7035 g cm⁻³ (ref. 43) and simply modifies the intercept of previous equations⁴⁴. Analysis of buoyant density was carried out by the method of Vinograd and Hearst⁴⁴ using *Micrococcus lysodeikticus* DNA as a marker. The θ value for the latter was found to be 1.7271 g cm⁻³ relative to SV40 DNA. Buoyant densities were determined in a Beckman model E analytical ultracentrifuge. All experiments were carried out at 40,000 r.p.m. and 20°C using a solvent consisting of 10 mM Tris, 1 mM Na₂EDTA (pH 8.0) and the appropriate amount of CsCl (Metallgesellschaft). Cells were scanned in conditions which yielded a magnification factor of ~19. The chart recording was measured with a precision ruler (Keuffel and Esser). The total uncertainty in the measurement of the marker and sample peaks on the scanner trace was ± 0.04 cm, corresponding to an experimental precision of θ of ± 0.0003 g cm⁻³, equivalent to ± 5.4 CMC molecules per molecule of DNA. The determination of θ is the major source of experimental error as transcription studies have shown that there is no greater than 2% difference between identical repeat reactions.

11 CMC molecules per SV40 DNA, thus excluding the possibility that CMC modification removes sufficient superhelical turns to reduce the torsional forces that may be responsible for structural alterations in promoters. Several models could explain why *E. coli* RNA polymerase binding sites are highly selective targets for CMC modification: (1) Supercoiling produces open promoters of single-strand character, allowing enhanced binding of RNA polymerase. This would explain the extreme sensitivity of promoters to CMC modification, as reactivity occurs at the imino sites of thymine and guanine and requires unpaired bases; (2) The promoter regions are made highly reactive due to structural transitions that are not presently understood; (3) CMC reactivity that occurs by either of the above models may be localised at sites adjacent to the promoter and perturb structural transitions that a promoter must undergo for appropriate recognition by RNA polymerase. This action at a distance or telestability²³ may be due to the CMC-induced disruption of "barriers, sinks or sources of sequence-specific conformational preference"²⁴ that may be needed to produce the recognition site for RNA polymerase.

Support for the localisation of altered secondary structure comes from a very recent study of Shishido²⁵ using *S*₁ endonuclease to map preferential unpaired or weak duplex regions in SV40. The cleavage sites were found at 0.15, 0.28, 0.38, 0.39, 0.44, 0.56, 0.89 and 0.98 SV40 map units. The average distance of an RNA polymerase molecule from a respective *S*₁ site is 0.032 map units (167 base pairs). Given an

uncertainty of ± 0.02 map units (103 base pairs) for the assignment of RNA polymerase binding sites by electron microscopy leads us to conclude that six of the *S*₁ sites are in close proximity to RNA polymerase binding regions.

If a promoter constitutes approximately 60 base pairs²⁶ one would anticipate that CMC modification would produce a differential effect on binding dependent on its location, that is directly in the promoter sequence or at adjacent sites upstream or downstream. This would eliminate binding in some cases but only weaken binding in others. However, if initiation is dependent on a precise RNA polymerase-DNA interaction one would anticipate a more drastic effect on transcription. This seems to be the case (Table 1) and is supported by our previous observation that rifampin sensitivity ($2-8 \mu\text{g ml}^{-1}$) is not seen until SV40 FI DNA is modified with CMC¹⁸. Hence one CMC per promoter produces a 37% loss of binding which represents an 88% decrease in transcription in the presence of $4 \mu\text{g ml}^{-1}$ rifampin.

All the RNA polymerase sites appear to be equally sensitive (Fig. 5) towards initial CMC modification. However, it has been shown that if CMC reactivity is allowed to continue for 24 h the reagent is redistributed as follows¹⁹: 80 CMC molecules mapping between 0.65 and 0.43 map units; 31 CMC molecules mapping between 0.32 and 0.17 map units; 10 CMC molecules mapping between 0.17 and 0.10 map units. This data is based on a low specific activity ¹⁴C-CMC probe available at that time that would not detect less than 5 CMC molecules per restriction fragment.

Thus it seems that supercoiling has the capacity to generate well separated localised sites. However, further chemically induced melting seems to coalesce open regions as a function of A-T content²⁷. Although the molecular details need to be resolved for a full understanding of the structural features of DNA that may be varied and generated by supercoiling, it is striking that altered or open secondary structure can be generated at highly preferential sites. This has considerable implications for biological recognition and the modulation of gene expression. Additional evidence supporting the role of supercoiling in transcription has been recently obtained for promoters in plasmids²⁸. Also, some promoters located in linear DNA are sensitive to novobiocin which inhibits DNA gyrase²⁸.

The study of SV40 DNA provides a useful model for the comparative analyses of *in vivo* and *in vitro* functional regions of DNA with regard to the expression and regulation of gene products. SV40 FI DNA can be transcribed by *E. coli* RNA polymerase and the resulting cRNA produces T antigen immunoprecipitated products in a variety of *in vitro* translation systems²⁹⁻³². In addition, linked transcription-translation systems have yielded the major capsid protein VP1, using *E. coli* RNA polymerase³¹⁻³⁴.

Figure 6 is a map of the early and late regions of SV40 DNA showing the sites of *E. coli* RNA polymerase binding. Figure 6a shows that *E. coli* RNA polymerase can initiate at six sites from the early region. These correspond quite well with the mapping

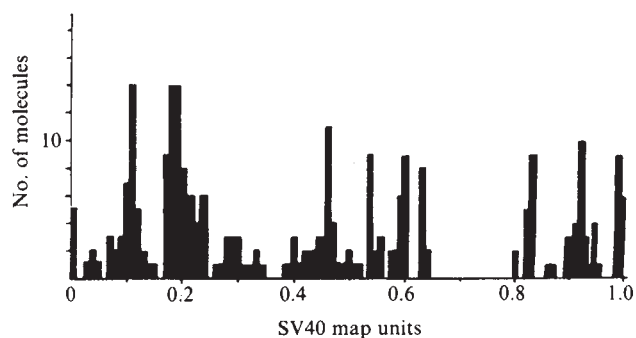
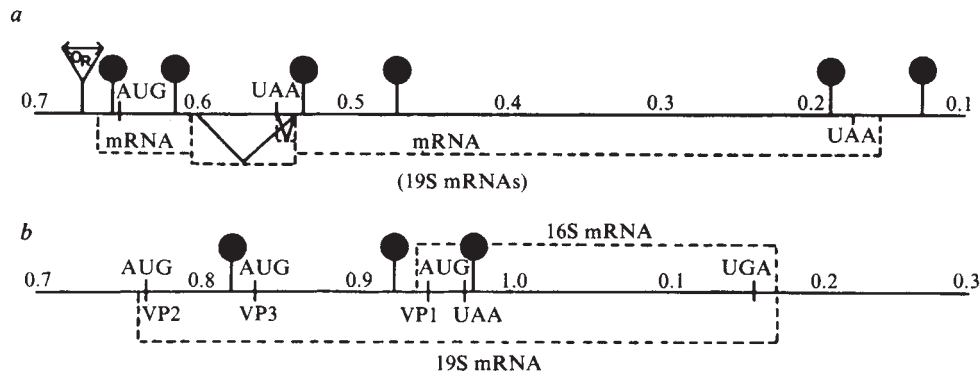


Fig. 5 Map location of *E. coli* RNA polymerase binding sites on 15 min CMC-modified SV40 FI DNA as determined by electron microscopy. Samples were prepared as described in Fig. 1.

Fig. 6 Location of the *E. coli* RNA polymerase molecules in the early (a) and late (b) regions of SV40 DNA. The respective regions of the SV40 DNA map are expressed in a linear form with relative lengths measured from the *EcoRI* position assigned 1.0 or 0.0. In both cases mRNA synthesis is expressed from left to right, 5' to 3'. The origin of replication is indicated by O_R and the cytoplasmic mRNAs are shown by dashed lines. a, The early region contains an intron sequence and two splices occur which are indicated by the large and small ∇ ^{33,45,46}. The two mRNAs that are produced consist of 0.65 to 0.55 and 0.65 to 0.60 fused respectively to the main body of the message 0.54 to 0.17. The larger mRNA, that is with the smaller spliced out regions, retains the UAA stop codon³³ shown at 0.547 and hence protein synthesis only produced a 17,000 molecular weight (MW) protein, little t antigen. The small mRNA, with the entire intron sequence spliced out produces large T antigen, with an approximate MW of 80,000. The UAA stop codon for this protein is shown at 0.175. b, The late region of SV40 coding for viral proteins VP1, VP2, and VP3. (ref. 33). The lower dashed line represents the 19S mRNA coding for VP2 and VP3 with AUG initiation codons at 0.767 and 0.835, respectively^{33,34}. These overlapping proteins are terminated by the stop codon UAA shown at 0.969 (ref. 33). The upper dashed line represents 16S mRNA coding for VP1 with initiation of protein synthesis from the AUG codon shown at 0.947 and termination at the UGA codon shown at 0.155 (refs 33, 34). The reading frame for VP1 (16S mRNA) is different than VP2 or VP3 (19S mRNA) and the UAA stop at 0.969 is not recognised³³. The respective AUG initiation and UAA or UGA termination codons have been determined from DNA and protein sequence data³³. No attempt has been made to show the variable 5' leader splicing that occurs from the region 0.721 to 0.760 which is attached to late 19S and 16S mRNAs³⁴. These may be involved in the control of the expression of VP2 and VP3.



of [γ -³²P] ATP initiations at 0.17, 0.45, 0.52, 0.61 and 0.65. The promotor at 0.65 would generate cRNA with translation stop signals³³, hence only the region from 0.65 to 0.55 can be translated. This would produce small t antigen. The promoters at 0.53 and 0.47 would produce cRNAs that could be translated into T antigen-like proteins. The two promoters between 0.20 and 0.10 should not produce any significant T antigen-like proteins. Figure 6b compares the *E. coli* RNA polymerase binding sites of the late region to *in vivo* functions. We observe that the enzyme binds before the appropriate initiation codons that would produce VP3 and VP1. The latter is the major capsid protein whereas VP2 and VP3 represent minor structural proteins^{24,33}.

It is striking that *E. coli* RNA polymerase recognises regions in the correct polarity proceeding AUG initiation codons for two out of three late gene products as well as the AUG initiation codon for T antigens. The other binding sites for the enzyme, except the 0.13 site, all occur before open reading frames with potential codons for initiation. With regard to primary sequence recognition sites, it has already been established that there is a promotor sequence for tight binding of *E. coli* RNA polymerase at 0.165 on the SV40 DNA map³⁵. Examination of the remaining SV40 DNA sequence (ref. 33 and P. T. Chan, unpublished data) reveals promoter sequences³⁶ that could fit with our electron microscopy RNA polymerase binding analysis with the exception of the 0.65 site. However, this site is very A-T rich³³ and may be easily opened to serve as a promoter. Many other potential promoter sequences are also available but not used, suggesting that some of the *E. coli* RNA polymerase recognition sites that are used share similar features with the eukaryotic RNA polymerase II. This is supported by a correspondence of five of six binding sites for the respective enzymes on polyoma DNA using electron microscopy mapping^{37,38}. A recent report indicates that *E. coli* RNA polymerase binds to the appropriate sites on adenovirus DNA for the production of some of the late mRNAs³⁹.

The evidence that supercoiling has a significant role in the modulation of gene expression in prokaryotic systems is quite compelling. The possibility that modulation of supercoiling occurs through the changes in binding of histone and non-histone proteins would be a plausible extension of these results. The finding that superhelical turns makes promoters highly preferential targets for carbodiimide modification opens new possibilities for exploring the structural transitions produced in DNA by the torsional stress of supercoiling.

We thank John Fitzpatrick for technical assistance, Peter Chan for discussions and Carolyn Harris for preparation of the

manuscript. This work was supported by grants CA17077 from the NCI and PCM 7808589 from the NSF. Patricia Hale was supported by National Research Service Award T32CA09128 from the NCI.

Received 30 June; accepted 13 November 1979.

- Lebowitz, J., Chaudhuri, A. K., Gonen, A. & Kito, G. *Nucleic Acids Res.* **4**, 1695-1712 (1977).
- Pritchard, A. E., Kowalski, D. & Laskowski, M. *J. biol. chem.* **252**, 8652-8659 (1977).
- Bauer, W. R. *Rev. Biophys. Bioeng.* **7**, 287-331 (1978).
- Bauer, W. R. & Vinograd, J. *J. molec. Biol.* **47**, 419-435 (1970).
- Davidson, J. *J. molec. Biol.* **66**, 307-309 (1972).
- Hsieh, T.-S. & Wang, J. C. *Biochemistry* **14**, 527-535 (1975).
- Bauer, W. R. & Vinograd, J. in *Basic Principles of Nucleic Acid Chemistry* Vol. 2 (eds Cantoni, G. L. & Davies, D. R.) 265-303 (Academic, New York, 1974).
- Hayashi, Y. & Hayashi, M. *Biochemistry* **10**, 4212-4218 (1971).
- Botchan, P., Wang, J. C. & Echols, H. *Proc. natn. Acad. Sci. U.S.A.* **70**, 3077-3081 (1973).
- Wang, J. C. *J. molec. Biol.* **89**, 783-801 (1974).
- Richardson, J. P. *Biochemistry* **13**, 3164-3169 (1974).
- Zimmer, S. G. & Millett, R. L. *Biochemistry* **14**, 300-306 (1975).
- Mandel, J. L. & Chambon, P. *Eur. J. Biochem.* **42**, 367-378 (1974).
- Flashner, M. S., Katopes, M. A. & Lebowitz, J. *Nucleic Acids Res.* **4**, 1713-1726 (1977).
- Hale, P. & Lebowitz, J. *J. Virol.* **25**, 298-304 (1978).
- Richardson, J. P. *J. molec. Biol.* **91**, 477-487 (1975).
- Botchan, P. *J. molec. Biol.* **105**, 161-176 (1976).
- Hale, P. & Lebowitz, J. *J. Virol.* **25**, 305-311 (1978).
- Chen, M., Lebowitz, J. & Salzman, N. P. *J. Virol.* **18**, 211-217 (1976).
- Lebowitz, P., Stern, R., Ghosh, P. K. & Weissman, S. M. *J. Virol.* **22**, 430-445 (1977).
- Lebowitz, J., Garon, C. F., Chen, M. & Salzman, N. P. *J. Virol.* **18**, 205-210 (1976).
- Bauer, W. R. & Vinograd, J. *Ann. N.Y. Acad. Sci.* **164**, 192-225 (1969).
- Wells, R. D. *et al. CRC Crit. Rev. Biochem.* **4**, 305-340 (1977).
- Selsing, E., Wells, R. D., Alden, C. J. & Arnott, S. *J. biol. Chem.* **254**, 5417-5422 (1979).
- Shishido, K. *Agric. biol. Chem.* **43**, 1093-1102 (1979).
- Gilbert, W. in *RNA Polymerase* (eds Losick, R. & Chamberlin, M.) 193-205 (Cold Spring Harbor Lab. Monogr. Ser., 1976).
- Mulder, C. & Delius, H. *Proc. natn. Acad. Sci. U.S.A.* **69**, 3215-3219 (1970).
- Yang, H.-L., Heller, K., Gelbert, M. & Zubay, G. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3304-3308 (1979).
- Prives, C., Gilboa, E., Revel, M. & Winocour, E. *Proc. natn. Acad. Sci. U.S.A.* **74**, 457-461 (1977).
- Paucha, E., Harvey, R. & Smith, A. E. *J. Virol.* **28**, 154-170 (1978).
- Roberts, B. E. *et al. Proc. natn. Acad. Sci. U.S.A.* **72**, 1922-1926 (1975).
- Greenblatt, J. F., Allet, B., Weil, R. & Ahmed-Zadeh, C. *J. molec. Biol.* **108**, 361-379 (1976).
- Reddy, V. B. *et al. Science* **200**, 494-502 (1978).
- Reddy, V. B., Ghosh, P. K., Lebowitz, P. & Weissman, S. M. *Nucleic Acids Res.* **5**, 4195-4213 (1978).
- Dhar, R., Weissman, S. M., Zain, B. S., Pan, J. & Lewis, A. M., Jr *Nucleic Acids Res.* **1**, 595-614 (1974).
- Scherer, G. E. F., Walkinshaw, M. D. & Arnott, S. *Nucleic Acids Res.* **5**, 3759-3773 (1978).
- Lescure, B., Chestier, A. & Yaniv, M. *J. molec. Biol.* **124**, 73-85 (1978).
- Lescure, B., Daguette, C. & Yaniv, M. *J. molec. Biol.* **124**, 87-95 (1978).
- DeLorbe, W. J. thesis, Univ. Iowa (1978).
- Gonzalez, N., Wigg, J. & Chamberlin, M. *Arch. Biochem. Biophys.* **182**, 404-408 (1977).
- Mangel, W. & Chamberlin, M. *J. biol. Chem.* **249**, 2995-3001 (1974).
- Kleinschmidt, A. K. in *Methods in Enzymology* Vol. XII B (eds Colowick, S. P. & Kaplan, N. L.) 361-377 (Academic, New York, 1968).
- Szybalski, W. & Szybalski, E. H. in *Procedures in Nucleic Acid Research* Vol. 2 (eds Cantoni, G. L. & Davies, D. R.) 311-354 (Harper and Row, New York, 1971).
- Vinograd, J. & Hearst, J. *Prog. Chem. Org. nat. Prod.* Vol. 20 (ed. Zechmeister, J.) 373-422 (Springer, Vienna, 1962).
- Crawford, L. V. *et al. Proc. natn. Acad. Sci. U.S.A.* **75**, 117-121 (1978).
- Berk, A. J. & Sharp, P. A. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1274-1278 (1978).

BOOK REVIEWS

Intelligence for beginners?

Donald D. Dorfman

EYSENCK'S new book on intelligence is his first to appear since Leon Kamin reported gross irregularities in Sir Cyril Burt's data. The main purpose of his book is to present "what to the author appears to be the paradigm towards which the research of the past 80 years converges" (page 6). The book is "intended for beginners" (page 6). An author of a work intended for beginners has an obligation to give a balanced and scrupulous description and analysis of the historical and scientific data.

On page one of his book, Eysenck refers the reader — presumably a beginner — to his appendix A which "gives a short account of the facts of the Burt affair". He implies in that appendix that Arthur Jensen was the first to discover anomalies in Burt's data. Leon Kamin — the actual discoverer — is never mentioned in that historical account.

In his first chapter, "Intelligence: The Development of a Concept", Eysenck concludes that culture-bound intelligence tests are admissible for pupil and student selection at school and university because "we are often justified in assuming considerable uniformity in cultural background among candidates" (page 23). Eysenck gives no justification for his cultural-uniformity hypothesis.

His review of the evidence for a correlation between IQ and amplitude and latency of evoked cortical potentials is misleading. The figures that he displays in support of a strong correlation are taken from Ertl's early work which even Eysenck admits "suffered from technical and methodological deficiencies" (page 50). He then asserts that Shucard and Horn have also obtained "quite sizable correlations between AEP's [averaged evoked potentials] and IQ" (page 50). In fact, those investigators reported a correlation of only +0.24 for fluid intelligence and an absence of correlation for crystallized intelligence in the article cited by Eysenck. He also presents data from "our own laboratories" collected in about 1973. No details are given and no reference is made to any relevant publications in scholarly journals.

Eysenck's introductory text is also interspersed with political comments. For instance, in his discussion of IQ and job success, he tells us that "in our type of society there is little evidence of political

The Structure and Measurement of Intelligence. By Hans J. Eysenck. Pp.253. (Springer: Berlin, Heidelberg and New York, 1979.) DM 44, \$24.20.

and ideological interference with employment and promotion (except perhaps through union activity)" (page 88). In his final chapter "Intelligence and Society", Eysenck reports only one example of abuse of IQ tests for job selection: "it has been shown that on occasion white trade unionists in the U.S.A. have insisted on the use of irrelevant IQ tests for job selection in order to keep out black applicants" (page 221).

"Does IQ Measure Intelligence?" is the title of Eysenck's chapter on the validity of the IQ test. "We shall consider some outstanding studies" (page 79) in this chapter, he writes. One of those "outstanding studies" presented was the investigation performed during World War I by American Army psychologists with the Army Alpha, a group test of intelligence. Eysenck gives some of their evidence for the validity of the intelligence test, for example, their finding that illiterate enlisted men performed more poorly on the Army Alpha than literate enlisted men. Eysenck, however, neglects to present any of the items from that IQ test. Here are a few: (a) "Harvard University is in ____"; (b) "Yale University is at ____". Here are a few more: (a) "Why is tennis good exercise?"; (b) "Lob is a term used in ____"; and (c) "a battle in racket very tennis useful is", which is to be rearranged into a sentence and then answered "true" or "false." The Army Alpha was the first IQ test used to classify national and racial groups according to native intelligence. "The great and universally agreed success of these tests caused many other countries to adopt them in later years and presents another external validation criterion for IQ tests as measures of intelligence" (page 83) Eysenck writes.

The book also contains three chapters jointly authored by Fulker and Eysenck: a chapter on heredity, one on environment and a final chapter on socio-economic status. Their evaluation of the evidence supporting an environmentalist position is quite good overall. Their criticisms of the Heber study are reasonable and their presentation of Zajonc's confluence model

is clear and fair. On the other hand, Fulker and Eysenck's arguments for the genetic position are in no way convincing and often misleading.

In his introductory chapter, Eysenck states (page 1): "I have tried to rewrite the relevant chapters in the history of the intelligence testing movement without including Burt's now doubtful data". Whereas Eysenck may have excluded Burt's questionable data from his historical accounts, Fulker and Eysenck present and make rather extensive use of Burt's fabricated numbers in their biometrical analyses. For example, in their analyses of IQ correlations for monozygotic twins reared apart, they include Burt's group-test correlation of 0.77, asserting that "the main criticism [of Burt's data] concerns the individual test scores" (page 109). That statement is false. Indeed, the main criticism concerns his group-test correlation: it remained constant at 0.771 from 1955 to 1966 in spite of large increases in sample size. Fulker and Eysenck also include Burt's questionable IQ correlations in their analyses of data for siblings reared apart (Table 5.4) and for unrelated children reared in the same home (Table 5.5). In their biometrical analyses of educational achievement, they use Burt's correlations for monozygotic twins reared together, for dizygotic twins reared together and for monozygotic twins reared apart (Table 5.14). Their biometrical exercises lead them to conclude that genotype-environment interaction and genotype-environment co-variance are small enough to be ignored. That conclusion contradicts the results of the statistical exercises presented in their final chapter on socio-economic status. The conclusions are as varied as the models.

Near the end of the book, there is an unnumbered chapter bearing the title "Eysenck and the Splitting of the IQ". A careful study of the primary sources reveals that if the IQ was split, it was split by Furneaux. In fact, Furneaux's first publication on this matter was apparently in *Nature* in 1952 (170, 37-38).

Summing up, Eysenck's new book is inappropriate for beginners, but rather entertaining for experts. □

Donald D. Dorfman is Professor of Psychology in the Department of Psychology at the University of Iowa, Iowa City, Iowa.

Comprehensive palaeontology

J.C.W. Cope

The Encyclopedia of Paleontology. Edited by R.W. Fairbridge and D. Jabolonski. Pp.912. (Dowden, Hutchinson and Ross: New York. Distributed by Academic Press: New York and London, 1979.) \$90, £52.50.

THE appearance of a major reference work on palaeontology within the compass of a single volume is indeed welcome. *The Encyclopedia of Paleontology* contains entries from over 120 palaeontologists, mainly from the USA, but with expertise from Europe, Australasia and Canada contributing some 30% of the information. Most authors have contributed only one entry, others have supplied as many as six.

The subject entries are very wide-ranging and cover in considerable detail each group of plant and animal fossils in alphabetical order. In addition to the taxonomic treatment there is a long list of other subjects covered such as biometrics, computer applications in palaeontology, palaeoecology, population dynamics and so on, in which the state of the art has clear exposition. These two categories of entries are complemented by longer discursive entries on such topics as the history of palaeontology (divided into three separate entries, historically), evolution, extinction and other subjects that lend themselves to such treatment.

The entries are well organized and provide a wealth of accurate information; each has cross-references to other entries and, where appropriate, to entries in other volumes of the *Encyclopedia of Earth Sciences*. The reader is directed to further reading by generally well selected and ample references. Occasionally one feels that the author of an entry has not been sufficiently selective in citation of these; as examples of over-zealous quotation of references "Paleoecology of Inland Aquatic Environments" has 110 references and "Radiolaria" almost 120. This surely is not helpful to the average reader seeking direction to further major references.

The references provide at least a general clue to the date of writing of the entry. Some seem to have been written about ten years ago, for although recent references are listed, no reference is made to them in the text. However, the entries in general are up to date; many articles were clearly written around 1976 and references up to 1978 occur commonly. Within the limits imposed by getting such a volume published this seems very reasonable. Where work is dated, it is largely a reflection of how rapidly some aspects of palaeontology have advanced recently.

On first glancing through the book I

wondered whether it had been mis-bound as there are no entries under 'N' — where were nannoplankton, nautiloids, nekton, neoteny? In fact these fears were largely unjustified as most of the above subjects appear elsewhere under other headings and are listed in the extensive index which, a quick calculation suggests, contains close on 5000 subject entries, many with several page references. The index must be criticized, though, in that it sometimes refers to words which are just mentioned in passing. Thus my search for 'neoteny' led to p.34 which merely stated that certain Cretaceous Amphibia were "pedomorphic or neotenic". The index entry for "pedomorphic" only quotes p.34 so that the reader is still none the wiser on what neoteny is, and both index entries are unnecessary. Checks revealed a number of similar superfluous entries. Misprints in the text have led to index entries; thus "Nautiloids" (*sic*) refers to a misprint on p.533 and is quoted by the index separately from "Nautiloids" and "Nautiloidea". "Conchostraca" get separate indexing from "Conchostracans", but the index entry "Cyclostomata" refers to both the vertebrate and bryozoan orders of that name. Clearly the index was not compiled by a palaeontologist, and its value is consequently somewhat diminished.

Although the book is extremely comprehensive in its treatment, I was surprised that scant attention had been paid to the stratigraphical use of fossils. There is an apposite entry on

biostratigraphy by Peter Sylvester-Bradley (one of seven authors who have since died) — but that is all. There is no entry in the index under the word 'zone', and although some systematic entries do refer to stratigraphical value of particular fossil groups, to me this is a major omission.

The book is abundantly illustrated by both half-tone and line illustrations; the great majority of them are first class and contribute significantly to the volume. The high standard of the majority makes one aware of the number which fall short of this standard in some respects; poor reproduction mars several otherwise good figures, and widely disparate sizes and styles of lettering are distracting.

The price of \$90 (or £52.50) seems excessive for a single volume. At half this price the number of purchasers would probably be increased at least tenfold; but the considerable authority and wealth of information contained therein should compel all palaeontologists seriously to consider buying the book. For undergraduates it will provide valuable information and speedy reference to essential reading for essays and project work; for university libraries at least one copy is essential. Despite its faults this is a volume truly worthy of its title and one which is unlikely to be rivalled for a long time. □

J.C.W. Cope is Senior Lecturer in Geology at the University College of Swansea, UK.

Kinetics for chemists

C.F. Wells

Chemical Kinetics and Transport. By P.C. Jordan. Pp.340. (Plenum: New York and London, 1979.) £13.55.

As he states in his preface, and is implied in the title, the author adopts a new style in his treatment of kinetics. The first three chapters are concerned with the fundamentals of the transport of molecules: two deal with the kinetic theory of gases, covering the properties of a collection of gaseous molecules at equilibrium and the transport of molecules; but the third introduces an unusual note for a book on kinetics, a treatment of electrolytic conduction and diffusion. In Chapter 4, the experimental approach to kinetics begins with the determination of rate laws. This is followed by a treatment of stationary state mechanisms, photochemistry and non-stationary state mechanisms. The book is then completed by a consideration of single collisions, the theory of reaction rates and various models for collisional processes.

In its concentration on chemical kinetics from a transport viewpoint, especially in the earlier chapters, the book has a novel

approach. Moreover, much important material is reserved for the problems at the end of each chapter; undergraduate students attempting to follow their more conventional lecture courses with the aid of this book might well become confused. The latter state will probably be increased by what some will regard as a cavalier treatment of elementary rate laws. In Chapter 4, complex ideas are introduced too early: for example, before single rate laws are explored, the problems of the $H_2 + I_2$ and the $H_2 + Br_2$ reactions are stated, and we find a discussion of nanosecond flash photolysis; or, before a discussion of simple orders higher than one, we are plunged into opposed first order reactions. Indeed, in a footnote the reader is referred to a standard text to obtain a detailed understanding of the integrated rate laws. Perhaps this is the most unsatisfactory chapter from an undergraduate point of view, as there are also minor contributions to his confusion. For example, in the section on first order reactions, one finds a jumble of first order and opposed first order reactions; or, on p.90, it is stated that the *first* term in eqn 4.31 vanishes, whereas, in the balance of rates at equilibrium, the *first two* terms disappear. However, if the student survives this chapter, or is assisted

in grasping these basic ideas from other sources, he should be able to proceed readily with the rest of the book. It is, perhaps, unfortunate that some of the theories of reaction rates are not introduced earlier than Chapter 9, so that the ideas developed there could be used *en route*.

Finally, I must introduce what is perhaps a more personal note, although I am sure that others also involved with kinetics in

solution will echo my view. Most general books on kinetics concentrate on the gas phase; understandably so, as here individual molecular processes can be examined in detail. Yet rarely are the complexities of processes in solution dismissed so completely: these are merely referred to very briefly in one paragraph in Chapter 5.

Although I have obviously read this book with a critical eye, I realize, of course, that not everything can be included in a

book of this size. Moreover, I am sure that all who are interested in kinetics will find it stimulating, and this includes undergraduates. It certainly should find a place on library shelves. The format of the book is good and it is well produced; the price is not unreasonable for these days. □

C.F. Wells is Reader in Inorganic Chemistry at the University of Birmingham, UK, and has particular interest in the kinetics of inorganic reactions in solution.

Volcanoes revisited

P.E. Baker

Volcanology. By H. Williams and A. R. McBirney. Pp.397. (Freeman, Cooper & Co.: San Francisco; Blackwell Scientific: Oxford, 1979.) £19.50, \$33.50.

THE authors preface their book with the modest statement that it "... does not purport to offer much that is new; it attempts rather, to summarize for the student and professional geologist current knowledge of volcanoes, particularly their behaviour and physical structure". Despite a superficial resemblance to earlier books on the subject of volcanoes (for example, Macdonald, 1972) and some fairly conventional chapter headings, the approach is essentially rather different. It is a textbook aimed at a specialist readership and largely devoid of dramatic accounts of volcanic eruptions and their consequences. It is written in a lucid and cautious style with every effort being made to adopt a quantitative approach. It attempts a rigorous explanation of volcanic processes and products in terms of established physical principles. There are more descriptive chapters, such as those on "Principal kinds of eruptions" and "Cones, domes and shields", which contain innumerable examples drawn from the authors' wide experience.

They pick their way delicately through the subject, rejecting untenable generalizations and sometimes reconsidering discarded theories. They remind us that the evidence for extraordinary amounts of igneous activity along the axes of oceanic ridge crests is largely circumstantial, and that potassium does not invariably increase towards the interior of continents. Perhaps, also, there is some truth in Daly's theory of the 'vitreous substratum' source for basalt.

At the end of the opening chapter the authors state that, "By far the greatest impetus for studies of igneous processes has come from a growing recognition of the central role they play in nearly all modern tectonic theories". Maps of active volcanoes and plate boundaries appear on the inside covers of the book but can scarcely be regarded as reflecting the

dominant theme of the text. There are references to sea-floor spreading, mid-plate volcanism and subduction of oceanic lithosphere, but the authors appear anxious not to overemphasize the link between volcanism and plate tectonics.

The scope of the book has been carefully restricted and there is, for example, very little discussion of the role of volcanism in the evolution of the Earth, or of volcanism and mineralization. Volcanic activity in the rest of the Solar System is also excluded apart from occasional references and photographs of the martian volcano Olympus Mons and a NASA picture of

volcanic activity on Io.

The book is generously illustrated with many first-class photographs and the occasional indifferent one (for example, Fig. 7-11). The figures are drawn in somewhat varied styles but are not always to the standard one might have expected. Nevertheless, in its description and interpretation of volcanic phenomena this is an excellent book and the authors have certainly succeeded in the task they set themselves. □

P. E. Baker is Professor of Geology at the University of Nottingham, UK.

Methods for studying evolution

R.A. Fisher

Biochemical Systematics and Evolution. By A. Ferguson. Pp.194. (Blackie: Glasgow, 1979.) £14.50.

THE popular image of a taxonomist is of a cantankerous anatomist interminably arguing with an equally cantankerous opponent over some apparently trivial detail. These days, it seems, anatomists are out and biochemical geneticists are in. Cantankerousness and argument are, of course, always with us, but perhaps for those who wish to understand rather than stake their reputations, biochemical systematics can throw some new and useful light on the processes of evolution.

Evolution depends on mutations in DNA, and the basic idea behind biochemical taxonomy is to detect the mutations themselves rather than their effects at several removes. Events overtake us with distressing rapidity these days and it is unfortunate for Dr Ferguson in his book *Biochemical Systematics and Evolution* that he missed the opportunity of mentioning the new technique of restriction mapping, since this allows one to examine the actual gene at first hand. But it is neither the easiest nor the cheapest of methods and for some time to come the electrophoretic comparison of proteins will certainly remain the method of convenience, if not of choice. It involves the

investigation of the gene product rather than the gene, but it is much closer to the ideal than the morphometric comparisons of anatomists.

Electrophoresis is a well established and well described technique and much has been published on the interpretation of protein separations and zymograms. It is a pity, therefore, that, having stated that the aim of this book "is to provide the reader with the necessary background to the principles and practice of electrophoresis", the author does not make a better job of it. If this really was his aim, the sections on methods are too brief and in parts misleading. If one were, for example, to try electrophoresis according to the diagram on page 36, one would fail. Electrical contact between the gel and the electrodes is apparently prevented by a plastic film. The chapter which deals with the interpretation of isozyme patterns is again too brief and sometimes inaccurate. In particular, the origins of isozymes could have been more clearly explained.

The second half of the book deals with the interpretation of data, and here Dr Ferguson seems to be on much more familiar territory. He deals concisely and clearly with the various ways of measuring evolutionary distance and the mathematical examples are well explained and easy to follow. It is this part of the book that will really be of use to the audience for which it is intended, "advanced undergraduate and postgraduate students". □

R.A. Fisher is Head of the Comparative Genetics Section in the Nuffield Laboratories of Comparative Medicine at the London Zoo, UK.

Metals and gases

J. M. Thomas

Chemistry of the Metal-Gas Interface. By M.W. Roberts and C.S. McKee. Pp.594. Clarendon Press/Oxford University Press: Oxford, 1979.) £26.

MUCH thought has gone into the planning and writing of this commendable book. Where does one begin in the ordering of a topic that is so expansive and multifaceted? Short summaries, however refreshing, tend to distort the truth. Usually they do not faithfully reflect the vast ramifying nature of a complex and important subject. But comprehensive and detailed surveys dull the mind, provoke the thought that "he who increases knowledge increases sorrow" and inevitably tend to bury important signals in oceans of noise.

A book that runs to nearly 600 pages may at first sight appear to belong to the second of these two categories. Not at all. It is so expertly organized and sub-divided into compact but cleanly dove-tailed components that the 13 chapters make for informative and pleasant reading. The first two chapters neatly summarize the basic concepts of the surface chemistry and crystallography of solids. The next two present authoritative accounts of low energy electron diffraction (LEED) for surface characterization, and the various branches of electron spectroscopy (X-ray or UV induced photoelectron spectroscopy, Auger electron spectroscopy) for the characterization of the energy levels of adsorbates and energy levels or bands of adsorbents. In both these topics, the authors, especially Roberts, have made several original contributions. A further two chapters continue on the theme of surface identification and characterization of the adsorbed linkages; one is devoted to the direct spectroscopic approach (infrared and Raman), the other to a short survey of the surface kinetics involved in such processes as molecular beam scattering. There is also a useful account of the kinetic formulations required to rationalize observed rates of adsorption and desorption as measured by, for example, dynamic mass spectrometry.

Following a short illuminating treatment of physical adsorption at 'clean' metal surfaces and the rudiments of surface diffusion, the book devotes four chapters (nearly 200 pages) to gas reactions at the (predominantly) single crystal faces of various transition and refractory metals — for example W(211), Pt(111), Cu(100), and Ni(111). Almost exclusive attention is paid to the behaviour of the diatomic gases O₂, H₂, CO and N₂; metal adsorption on metals is considered; and the particular activity of stepped surfaces (first clearly identified by Somorjai) is reviewed. This section of the book succeeds admirably in placing recent developments in the context

of earlier endeavour, one of the declared aims of the authors. This section is likely to serve as a valuable quarry for all those interested in retrieving critically digested and reliable information of metal interaction with simple gases.

The final chapter on the impact of new experimental developments on heterogeneous catalysis is a little disappointing. Greater emphasis could have been given to the comparability of certain processes, such as the activation of CO and O₂ in homogeneous and heterogeneous systems. The recent work of Muetterties *et al.* (1979) on metal cluster compounds as models for heterogeneous catalysts, and analogous studies by Bradshaw *et al.* (1979) on Ru₃(CO)₁₂, draw attention to the kinship between the exterior of metal clusters on the one hand and exposed metal surfaces — for example Ru(001) — on the other. Similarities in densities of state alone (as exist between tungsten carbide and platinum at their Fermi levels) are not likely to be the sole cause for the observed,

uncanny similarity in catalytic activity of these materials as reported by Levy and Boudart in 1973.

One may point to a few errors of omission: no mention is made of EXAFS nor of the remarkable studies made (by Yagi and others) on reflection electron diffraction/microscopy of clean and contaminated 'elemental' surfaces. EXAFS in particular, with its ability to yield bond distances, coordination numbers and mean square fluctuations of neighbouring atoms, is already making a valuable contribution to our understanding of solid surfaces. But these are very recent advances. The authors had to call a halt somewhere. How many books in this field published less than a decade ago discussed XPS, UPS, AES and several other techniques which are so impressively paraded in this volume? □

J. M. Thomas is Professor and Head of Department of the Department of Physical Chemistry, University of Cambridge, UK.

Anion determination

H.M.N.H. Irving

Handbook of Anion Determination. By W.J. Williams. Pp.630. (Butterworths: London, Boston, Sydney and Toronto, 1979.) \$45.

ANALYSTS familiar with Sandell's *Colorimetric Determination of Trace Metals* have long felt the need for an equally authoritative text on anions. Faced with the need to determine a specific anion (and the present book deals with some 75 different ones), we must select one from among the many procedures which may have been published after due consideration of the accuracy and precision required, the length and cost of a determination, whether a few or a large number of samples will be presented, and the apparatus and the operator's skill needed.

The author deals in order with 36 common anions from acetate to vanadate, then with 14 halogen anions, 14 phosphorus oxyanions, and finally 11 sulphur anions. For each anion there is a well referenced review of separation and concentration procedures (which include distillation, liquid-liquid extraction, chromatography and ion-exchange, precipitation, adsorption and special methods as appropriate). Then follows a review of determinations by gravimetric, titrimetric, spectrophotometric and electroanalytical methods emphasizing techniques such as infrared spectroscopy, the use of ion-selective electrodes, atomic absorption and fluorescence, radiometry and miscel-

laneous methods. The reader must usually make his own choice of procedure but in many cases experimental details of selected determinations are given in full. Where reference is made to special apparatus there are excellent line drawings and great pains have been taken to give the structures of the many organic reagents used; their lay-out is, however, inconsistent in that methylene blue is shown as a sulphonium cation on page 550 and as an ammonium salt on page 576, and the formulae for its fluoroborate complex on page 34 is patently incorrect as must be those of the ternary zirconium-calcein blue-F⁻ complex on page 369 and nitron on page 123. Dyes derived from triphenyl methane may sometimes appear with a carbonium ion, a quinonoid structure, or a lactone ring: ionization of a salt is sometimes explicit but the older presentation, for example, CH₃C₆H₄SO NC1Na, is typical of other formulae in which salt formation must be understood.

But this is purely carping at points which will not diminish the great value of this book to practising analysts who have long had to rely on a miscellany of publications for material that is gathered together here in a single comprehensive book. I particularly liked the full and critical discussions of the determination of phosphorus as molybdo- and vanadophosphate, and the determination of halides, alone and in admixture. Despite the price, no analyst or teacher of inorganic chemistry can afford not to have this fine new undertaking within his easy reach. □

H.M.N.H. Irving, formerly Professor of Structural and Inorganic Chemistry at Leeds University, is now Professor of Analytical Science at the University of Cape Town.